The Relation between Physical Properties of Galaxies and their Environmental Geometry in the Sloan Digital Sky Survey

Markus B. Huber

München 2015

Ludwig-Maximilians-Universitat Munchen

The Relation between Physical Properties of Galaxies and their Environmental Geometry in the Sloan Digital Sky Survey

Dissertation

Executed at the
Max-Planck-Institut fur extraterrestrische Physik

Advisor: Prof. Dr. Gregor Morfill

Author: Markus B. Huber
Bruckmuhl

Submitted in May 2015

Erstgutachter: Prof. Dr. Dr. Gregor Morfill

Zweitgutachter: Prof. Dr. Andreas Burkert

Tag der mündlichen Prüfung: 16.12.2015

Zusammenfassung (German)

Die großraumige Galaxienverteilung besteht aus Galaxien in konzentrierten Galaxienhaufen und entlang Filament-ahnlicher Strukturen, die sich am Rande von fast leeren Gebieten im Universum (Voids) befinden. Das jeweilige lokale Umfeld fuhrt zu unterschiedlichen Entwicklungsverlaufen von Galaxien, die durch photometrische und spektroskopische Beobachtungsdaten charakterisiert werden. Um die unterschiedlichen Galaxienumgebungen in Beobachtungsdaten oder numerischen Simulationen zu klassifizieren, ist es ublich, die lokale Dichte jeder Galaxienumgebung zu schatzen, was z.B. zu der Morphologie-Dichte-Relation fuhrt. In dieser Arbeit wird zusatzlich die lokale Geometrie jeder Galaxie mit der Skalierungs-Index-Methode (SIM) geschatzt, die die Form der lokalen Galaxienverteilung charakterisiert, z.B. die vorkommenden Filamentstrukturen in manchen Galaxienregionen. Die neue Anwendung von SIM liefert Strukturinformationen uber die Galaxienumgebungen, die mit den photometrischen und spektroskopischen Galaxieneigenschaften korrelieren. Dazu wurde eine Volumen-limitierte Stichprobe ($0\ 05 < z < 0\ 1$) mit 93873 Galaxien aus dem Sloan Digital Sky Survey (SDSS) extrahiert und deren lokale Dichte und Geometrie fur verschiedene Radien geschatzt. Die Stabilitat der Geometrieschatzung, deren Klassifizierungsfehler und der Einfluss der Rotverschiebung wurden mit der Bootstrapmethode, mit Strukturprototypen und mit simulierten Galaxienverteilungen untersucht. Verschiede Galaxienhelligkeits-, Farb-, und Morphologieeigenschaften, sowie Indikatoren fur Sternentstehung wurden aus dem SDSS-Katalog extrahiert. Zusatzlich wurde aus diesen Daten die Sternentstehungsrate geschatzt. Die gefundenen Relationen weisen darauf hin, dass die physikalischen Galaxieneigenschaften nicht nur mit der lokalen Dichte korrelieren, sondern auch mit der von SIM geschatzten lokalen Geometrie, was z.B. zu einer entsprechenden Morphologie-Geometrie-Relation fuhrt. Basierend auf diesen Ergebnissen lasst sich feststellen, dass Galaxien in Filament-ahnlichen Regionen blauer sind, eine hohere Haufigkeit von spiralformigen Morphologien zeigen und eine hohere Sternentstehungsrate vorweisen, wenn sie mit der Population aus konzentrierten Galaxienhaufen verglichen werden. Der Vergleich mit Galaxien aus leeren Regionen im Universum zeigt, dass Galaxien in Filament-ahnlichen Regionen roter sind, eine geringere Anzahl von spiralformigen Morphologien haben und eine geringere Sternentstehungsrate besitzen. Selbst wenn der Einfluss der lokalen Dichte korrigiert wurde, konnten diese Abhangigkeiten der Galaxieneigenschaften von der lokalen Geometrie beobachtet werden. Diese Ergebnisse lassen darauf schließen, dass die lokale Geometrie die Beschreibung der Galaxienumgebungen durch die lokale Dichte erweitert. Die gefunden Zusammenhange zeigen, dass Galaxien selbst außerhalb von Gebieten mit hoher Dichte durch die lokale Umgebung in ihrer Entwicklung beeinflusst werden konnen. Dies lasst sich z.B. an Galaxien in Filament-ahnlichen Strukturen beobachten, die sich in ihrer Entwicklungsgeschichte und in ihren physikalischen Eigenschaften von anderen Galaxien unterscheiden.

Abstract

The large-scale galaxy distribution consists of galaxies concentrated in clusters and along filament-like structures, which are found on the edges of mainly empty, void-like regions in the Universe. These local environments can reflect different galaxy evolution histories that determine the observed photometric and spectroscopic properties. To classify different galaxy environments in surveys or numerical simulations, the local density of each galaxy can be estimated leading e.g. to the morphology-density relation. In this thesis, the local geometry of each galaxy is estimated using the Scaling Index Method (SIM), which characterizes the shape of local galaxy distributions, e.g. filament-like galaxy neighborhoods. This new SIM application results in structural information about the galaxy environments that is correlated with a set of photometric and spectroscopic galaxy properties. Specifically, a volume-limited sample ($0\ 05 < z < 0\ 1$) with 93873 galaxies was extracted from the Sloan Digital Sky Survey (SDSS) data and the local density and geometry were estimated for a set of radii. The stability of the geometry estimation, classification errors, and the influence of the redshift space distortions were investigated using bootstrap analysis, structure prototypes, and mock catalogues, respectively. A set of galaxy luminosities, colors, morphology indices, and indicators of star formation were extracted from the SDSS data and, in addition, the star formation rate was calculated for all galaxies. The resulting relations indicate that physical galaxy properties do not only correlate with the local density, but also with the local geometry estimated by SIM leading e.g. to a morphology-geometry relation. According to these findings, galaxies in filament-like regions are bluer, contain more late-type morphologies, and have a higher star formation than galaxies in clusters. Compared with galaxies in void-like regions, galaxies in filament-like regions are redder, contain more early-type morphologies, and have a lower star formation. After correcting the relations for the influence of the local density, the dependence of the galaxy properties on the local geometry was still observed. This indicates that the local geometry can extend the density-based description of the galaxy environments. The presented relations suggest that galaxies outside high-density regions can still be affected by the local environment during their evolution, e.g. in filament-like structures, resulting in different evolution histories and observable properties.

Contents

Chapter 1

Introduction

Observations of galaxy locations in the Universe have shown that luminous matter is not randomly distributed. Instead, galaxies are concentrated in clusters and along filament-like structures, which are found on the edge of mainly empty, void-like regions in space. A key question in modern cosmology is to understand the evolution of the large-scale galaxy distribution in the Universe and how the characteristics of the environment of galaxies affects their evolution and properties. For instance, it is well known that intrinsic physical properties of galaxies correlate with the local density of the environment (Hubble & Humason 1931). Modern galaxy surveys aim to expand this understanding and allow to test theories explaining the origin of large-scale structure and the physical properties of the embedded galaxies. To this end, sophisticated and novel descriptions of the observed data can provide additional relations between the galaxy properties and the environment. The application of the Scaling Index Method to analyze the galaxy distribution represents such a new approach and delivers an unique and local structure characterization.

1.1 Structure Formation in the ΛCDM Model

The standard cosmological model describes with a few parameters both the evolution of the Universe and astronomical observations on a wide range of scales from solar systems to the distribution of galaxy clusters. In this power-law Λ cold dark matter (ΛCDM) model the Universe is spatially flat, homogenous and isotropic on large scales. Its energy density is composed of a small fraction of radiation and neutrinos, 4% ordinary baryonic matter like electrons, protons, and neutrons, 23% non-baryonic dark matter and 73% dark energy represented by Einstein s cosmological constant Λ. Galaxies and other cosmic large-scale structures grew gravitationally from tiny primordial fluctuations that were adiabatic, nearly scale-invariant Gaussian random fluctuations following a power-law spectrum. The ΛCDM model fits a wide range from of astronomical data like the cosmic microwave background measured by the Wilkinson Microwave Anisotropy Probe (WMAP) and the large-scale distribution of matter and galaxies as mapped by the Sloan Digital Sky Survey (SDSS) and the Two Degree Galaxy Redshift Survey (2dFGRS). A minimal model with six cosmological parameters is sufficient to predict the statistical properties of those surveys: the density of matter, the density of atoms, the expansion rate of the Universe, the amplitude of the primordial fluctuations, their scale dependence, and the optical depth of the Universe (Spergel et al. 2003, 2007; Tegmark et al. 2004a; Komatsu et al. 2010).

The ΛCDM model uses a general description for the dynamics of space-time together with a metric to measure distances that both can be derived from Einstein s field equations, which link the energy-impuls-tensor with the geometric properties of space-time (e.g. Hawking & Ellis 1973, Chap. 3). A special case of this approach is the *Robertson-Walker-Metric*, which follows directly from the *Cosmological Principle*, the assumption that space is homogenous and isotropic on large scales. The expansion of the Universe can then be described by a scale factor $a(t)$,

which measures the distance of two comoving points. This leads to the *Friedmann Equations* that express the evolution of the scale factor $a(t)$ as functions of the matter and radiation density ρ, the pressure p, and the cosmological constant Λ.

Initial conditions for observed structures were established in the very early Universe where cosmic inflation led to a homogenous, isotropic, and flat space (Guth 1981, Albrecht & Steinhardt 1982, Linde 1983). During the hot, radiation dominated epoch, the Universe underwent an accelerated expansion approximately 10^{-34} s after the Big Bang. This exponential expansion inflated small, causal connected regions and flattened the space-time geometry. Small quantum fluctuations in the relativistic gas were magnified to macroscopic density perturbations and gravitational waves. A further success of the Inflation theory is the accurate prediction of the nearly scale-invariant spectrum of the Gaussian density fluctuation. However, a complete scale invariant power spectrum $P(k)$ of these fluctuations, the so-called *Harrison-Zeldovich-Spectrum*, $P(k)$ k, is not the best fit to WMAP data (Spergel et al. 2007). These small primordial temperature fluctuations are the foundation for all cosmological structures that are imprinted in the cosmic microwave background and provide the seeds for the largest structures. After the reheating of the primordial plasma at the end of inflation, the plasma cooled down to the point of the big bang nucleosynthesis were the first nuclei of deuterium, helium and lithium were created. During this epoch, the dynamic of the cosmic scale factor $a(t)$ can be described by a radiation-dominated Friedmann Model with the adiabatic expansion $a(t) \sim t^{1\ 2}$; the cosmic temperature of the relativistic gas decreased with T a^{-1}.

As the Universe expanded and cooled further, particles became non-relativistic, the first nuclei formed from free protons and neutrons at a temperature $T \sim 10^9$ K, and the Universe became matter dominated. At the end of this period of recombination marked by a redshift of $z_{\rm rec} \sim 1000$, the nuclei had bound all free electrons, which suppressed the Thompson scattering with photons. Matter and photons were decoupled and the Universe became transparent. The temperature of the photons decreased further with T a^{-1} and can be observed as today s cosmic microwave background, the surface of last scattering, that follows a Planck spectrum at a temperature $T_0 =$ 2 728 K. Although the amplitude of cosmic structures did not grow substantially during this epoch, the perturbations, which entered the horizon, oscillated sinusoidally resulting in the baryon acoustic oscillations (Eisenstein et al. 2005) and in the cosmic microwave anisotropy. Matter density fluctuations decoupled from the overall expansion of the Universe; instead, over-dense regions attracted the surrounding matter due to gravitational forces.

Dark matter began to collapse into a complex network of dark matter halos well before ordinary matter for which the gravitational Jeans instability (Jeans 1902) is opposed by radiation pressure. For cold dark matter only the growth of small-scale fluctuations is prevented in the radiation dominated Universe due to the small free-streaming length much smaller than a galaxy-scale perturbation. The existing dark matter perturbations of different scales evolved independently and as the Hubble radius grew, larger density perturbations were causally connected and, during matter domination, all perturbations grew through gravitational clustering. After the decoupling of photons and electrons, luminous baryonic matter followed the gravitational clustering and mirrored the evolution of the dark matter. The linear power spectrum of this matter distribution has been estimated by galaxy surveys (Tegmark et al. 2004b) and the Lyman-α forest (Seljak et al. 2005). On the largest scales, the galaxies and quasars are expected to follow the density field of dark matter.

The structure evolution in the Universe began with dark matter that clumped immediately after the Big Bang forming spherical halos. The baryons, in contrast, were initially kept from clumping by their interactions with one another and with radiation and remained in a hot, gaseous phase. As the Universe expanded, this gas cooled and the baryons were able to pack themselves together. The first stars and galaxies coalesced out of this cooled gas in the centers of the dark matter halos a few hundred million years after the big bang (Lehnert et al. 2010). The first structures were small, low-mass dark matter halos, which merged with one another with a high probability to form larger-mass systems, a bottom-up process.

As long as the density deviations were small, the dark matter can be treated as a pressure-less

fluid that can be modeled by linear equations. When the local densities became substantially denser than the mean density of the Universe, those approximations are invalid, since the dark matter starts to form caustics in which particles trajectories cross and form orbits. More sophisticated approaches like the Zeldovich approximation (Zeldovich 1970) or the Press-Schechter formalism (Press & Schechter 1974) model the gravitational clustering further in the non-linear regime to a certain extend and provide useful predictions. However, the dynamics are best understood using N-body simulations suggesting that matter condenses in filaments and halos between large voids forming a web-like structure of galaxy groups and clusters (White et al. 1987; Springel et al. 2005).

1.2 Galaxy Evolution

During the evolution of the density fluctuations, gravity has collected the dark matter in concentrations in dark matter haloes of different scales. In larger dark matter haloes, baryons are dense enough to radiate away enough energy to collapse into galaxies and stars. The most luminous galaxies are formed in the most massive haloes, preferentially in regions of the highest local mass density. Less luminous galaxies are developed in the less massive haloes that are found in regions with lower local densities, in ridges of matter running between denser regions, forming a cosmic web of filaments and sheets. An example for this large-scale structure of the galaxy distribution can be observed in our immediate extragalactic neighborhood, the so-called Local Sheet (Peebles & Nusser 2010).

Most galaxies in the Universe are located in groups and clusters of galaxies, which in turn reside in larger systems - in superclusters of galaxies or in filaments crossing under-dense regions between superclusters. The largest and richest superclusters are the largest coherent systems in the Universe with characteristic dimensions of up to $100\,h^{-1}$Mpc. According to the Cold Dark Matter model, groups and clusters of galaxies form hierarchically through the merging of smaller systems (Knebe & Muller 2000). The timescale of the evolution of groups depends on their global environment (Tempel et al. 2009). As a result, the properties of groups depend on the environment in which they are embedded: richer and more luminous groups are located in a higher-density environment than smaller, less luminous groups (Einasto et al. 2003, 2005; Berlind et al. 2006). Hence, an understanding of the properties and evolutionary state of groups and clusters of galaxies in different environments is important for the study of the galaxy distribution.

The present-day dynamical state of clusters of galaxies depends on their formation history and galaxy interactions like merging affect the properties of galaxies in clusters (Binney & Tremaine 2008). The well-known morphology density relation states that early type, red galaxies are located in clusters (in the central areas) while late type, blue galaxies can preferentially be found outside of rich clusters, or in the outskirts of clusters (Butcher & Oemler 1978; Dressler 1980; Einasto & Einasto 1987). An old question is whether the properties of galaxies depend on the cluster-centric radius or on the local density of galaxies in clusters, or on both (Whitmore & Gilmore 1991; Huertas-Company et al. 2009; Park & Choi 2009; Park & Hwang 2009).

1.3 Relations between Environment and Galaxy Properties

Studies have shown that the physical properties of galaxies are correlated and that many of them depend on the stellar mass (Kauffmann et al. 2003a,b; Tremonti et al. 2004; Brinchmann et al. 2004). Massive galaxies consist of old stars, have a high mass-to-light ratio, low star formation rates, little dust attenuations, high concentrations and stellar mass densities, and they often host an active galactic nucleus (AGN, Miller et al. 2003). In contrast, low-mass galaxies have young stellar populations, low mass-to-light ratios, high star formation rates, low concentrations and surface mass densities, and they almost never contain an active nuclei. The gas-phase metallicity of emission-line galaxies is also related to the stellar mass (Tremonti et al. 2004); the characteristic

stellar mass of $\approx 3 \times 10^{10} M_{\odot}$ was found where many galaxy properties change rapidly.

Many other galaxy properties can depend on the surrounding environment as well. The galaxy luminosity function (e.g. Blanton et al. 2001, 2003b; Blanton & Moustakas 2009) is a fundamental tool for the interpretation of observations and describes the number of galaxies per unit volume as a function of the luminosity. In SDSS and 2dFRS data sets, the calculated luminosity functions depend strongly on the local density of the environment: galaxies are more luminous in rich clusters (De Propris et al. 2003) and marginally more luminous in groups (Martínez et al. 2002) as compared to galaxies in the field (Madgwick et al. 2002). It was also shown that void galaxies are fainter than wall galaxies (Hoyle et al. 2005), and that galaxies in filaments close to clusters are more luminous due to interactions that induce rapid star formation (Braglia et al. 2007).

Galaxies colors reflect their dominant star population and provides hints about the current star formation and star formation history. Two main groups can be identified in any galaxy surveys: red galaxies are dominated by old, metal-poor red giants, whereas blue galaxies consist of young, metal-rich stars. Although the observed galaxy colors correlate with the morphology (Humason 1936; Hubble 1936), colors can be studied for more distant galaxies in the SDSS (Strateva et al. 2001; Baldry et al. 2004; Balogh et al. 2004b). While the morphological types can be assigned with some certainty only for nearby well-resolved galaxies, it is not possible for the fainter and more distant galaxies. It has been shown that blue and red galaxies populate environments of different density (Hogg et al. 2003; Balogh et al. 2004b). The red population in galaxies is a strong function of the local density for fixed luminosities, increasing from a fraction of $\sim 20\%$ of the whole population in the lowest density environments up to $\sim 70\%$ in the highest; the blue population shows only a marginal dependence on the environment (Balogh et al. 2004b). Since the stellar population contains information about the star formation history and galaxy formation, the question arises what physical mechanisms led to this relation between the environment and the stellar population. A dominance of a red stellar population in galaxies can be explained by an inability to form new stars due to a lack of available interstellar gas, which has to be shared with neighbor galaxies e.g. in high-density regions like clusters. For galaxy clusters with redshifts $z > 0\ 2$, the fraction of blue galaxies is higher (Butcher & Oemler 1978, 1984).

The color of galaxies is closely related to their morphological appearance that can be qualitatively described with the well-known Hubble sequence (Hubble 1926). The standard morphological classification scheme mixes elements that depend on the structure of a galaxy, e.g. bulge-to-disc ratio, concentration ratio, and surface density, with elements reflecting the star formation history, like dust lanes and the spiral arms strength. Early-type galaxies show elliptical and lenticular morphologies, have a ceased star formation rate and consist of an old population of red giant stars. Late-type galaxies have spiral-like morphologies and are characterized by a high star forming rate consisting of mainly young blue stars, which dominate the observed photometric properties. These two main groups of galaxies dominate different environments according to the morphology-density relation (Hubble & Humason 1931; Dressler 1980; Postman & Geller 1984; Goto et al. 2003a). Early photometric surveys indicated that the fraction of early-type galaxies is higher in clusters of galaxies compared to low-density environments (Hubble 1926; Oemler 1974; Dressler 1980). The morphology-density relation appears to be an universal characteristic of galaxy populations (e.g. Postman & Geller 1984; Goto et al. 2003a; Smith et al. 2005). While a change of morphology requires drastic interactions like ram pressure stripping (Quilis, Moore & Bower 2000), the stellar population is mainly controlled by the supply of interstellar gas.

Since observations of the galaxy s morphology require the existence of luminous stars that indicate the baryonic matter distribution within the galaxy, morphology and star formation are correlated (e.g. Kennicutt 1998a; Christlein & Zabludoff 2005) and both are functions of the environment. Gómez et al. (2003) observed that the star formation rate (SFR) changes as a function of the local projected galaxy density using the SDSS Early Data Release. Especially strongly star-forming galaxies, show a strong decrease with increasing density. Consistent results were observed by Hashimoto et al. (1998), who investigated the influence of environment densities on the SFR in the Las Campanas Redshift Survey, and by Lewis et al. (2002), who studied the same relations in the 2dF galaxy sample. Both groups were able to identify a critical density and

radius, where the SFR of cluster galaxies changes from a specific SFR for field galaxies. At the turnaround radius $R_t \quad 5R_v$, with the cluster s viral radius R_v, the SFR becomes statistically identical to the field. These results are in agreement with numerical simulations suggesting that the main physical properties of the SFR is controlled by the amount of cool gas and the time since the last interactions with a larger halo. Furthermore, the observations of Gómez et al. (2003) are qualitatively consistent with the hierarchical models that structure formation can affect the SFR of galaxies even beyond the virial radius. Studies indicate that correlations between the star formation history of a galaxy and its environment extends to low densities and to large cluster-centric radii (Lewis et al. 2002; Gómez et al. 2003; Balogh et al. 2004a).

Many of the observed relationships between the galaxy properties and the environment can be explained with current models of galaxy formation and evolution. The properties depend either only on the formation history of the dark matter halos (nature) or on nurture processes, such as ram pressure and harassment. In the cold dark matter models, a galaxy-scale dark matter halo forms through a series of mergers of sub-galactic objects. The final mass is determined by the overdensity at the galaxy scale, which is statistically related to the density field at a larger scale. Although other galaxy properties will depend on the final mass, a high variation in the properties can be expected depending on different scenarios how the final mass is assembled. For instance, Gott & Thuan (1976) proposed that galaxy morphology is determined by the amount of interstellar gas left over at the maximum collapse of the protogalaxy; they claim that at this point, ellipticals had finished star formation. Not only the amount of stellar gas depends on the halo mass, but also the gas cooling processes. In dark matter halos of low mass, the infalling cold gas is never shock-heated and will collapse directly onto the disk. In high-mass halos, the infalling gas is first heated to the halo temperature by shocks. The gas remains pressure-supported and in quasi-static equilibrium while it cools by radiative processes (White & Rees 1978; White & Frenk 1991; Cole et al. 2000).

In addition to those initial conditions and sub-galactic scale evolution, the local environment can also affect the galaxy properties. In isolated areas of space, galaxies are able to accrete intergalactic gas and form a disk, while in high density regions like clusters, galaxy interaction are more likely, e.g. merging (Toomre & Toomre 1972; Kauffmann, White & Guiderdoni 1993; Mulchaey & Zabludoff 1999). Mergers operate most efficiently in galaxy groups or in the outskirts of rich clusters (Gnedin 2003). Although merging cross-sections are low in the viralised regions of clusters, galaxies can still be altered by the cumulative effect of many weak encounters (Moore et al. 1996). A mechanism called ram pressure stripping removes gas and shuts off star formation when a galaxy orbits a hot cluster halo (Gunn & Gott 1972; Abadi, Moore & Bower 1999; Quilis, Moore & Bower 2000). More processes that transform the morphology of spiral galaxies to early types include the so-called harassment, when high-speed encounters of galaxies with other halos typically in clusters cause impulsive heatings (Barnes & Hernquist 1991; Moore et al. 1996, 1998). Strangulation (Larson, Tinsley & Caldwell 1980; Balogh et al. 2000) can also transform morphology through a decline of star formation rate due to shut-off of the newly accreted gas when a galaxy enters a cluster or group environment and loses its hot gas reservoir. There is also observational and theoretical support for that tidal forces in clusters can transform the infalling spirals (Moss & Whittle 2000; Gnedin 2003). Another source of mutual influence are tidal interactions between close galaxies, which can induce bursts of star formation (Nikolic et al. 2004). It was shown in numerical simulations that weak bulges in spiral galaxies facilitate the creation of bars during such tidal interactions (Mihos & Hernquist 1996) and that tidally induced collisions of disk gas clouds are effective mechanism to create barred and S0 galaxies in rich clusters (Byrd & Valtonen 1990). These interactions create gas-poor elliptical galaxies in high-density regions, while low-density regions remain dominated by more gas-rich spiral galaxies with high star-formation rates. This evolution processes lead to many of the observed relationships, such as the morphology-density relation (Dressler 1980; Postman & Geller 1984).

1.4 Description of Large-scale Structure

Besides the observation of galaxy properties, the main objective of SDSS is to estimate statistical properties of the large-scale structure of the three-dimensional galaxy distribution. This distribution consists of galaxy clusters that are connected with filament-like and wall-like connections, which are found on the boundaries of large void-like regions. The commonly used statistical models for the main cosmological fields such as the density, velocity, and gravitational potential are based on Gaussian random fields (e.g. Bardeen et al. 1986), which are completely determined by their two-point correlation function (Totsuji & Kihara 1969; Peebles 1973, 1980) or power spectrum (Peacock 1999; Tegmark et al. 2004b; Cole et al. 2005). The main goal of spatial statistics in cosmology is to estimate these two functions that can be directly related to simulations. However, since both functions describe only second order statistics, they do not reveal any information about higher order or non-linear correlations in the galaxy distribution.

While these statistical measures are easily related to theoretical predictions and provide comprehensive information about the scale, amplitude, and even the nature of the deviations from a uniform distribution, they at best yield only suggestive and global statistical measures for these local structural patterns. To overcome this disadvantage, extension to this approach have been investigated. Higher order correlations were estimated with three-point and four-point correlation functions (Groth & Peebles 1977; Fry & Peebles 1978). The analysis of the density field in the Fourier space was expanded by estimating the bi-spectrum (Scoccimarro et al. 1998, 2001) and the Fourier space entropy (Coles & Chiang 2000), which can detect non-Gaussian signatures in the density field.

Different approaches have been proposed to extract more structural information from these complex patterns in the galaxy distribution. The topological analysis of the large-scale structure estimates the degree of connectivity of the galaxy distribution in the Universe. A first application of this approach is to describe the pattern in the galaxy distribution with the abundance of empty regions (Cressie 1993) using the void probability functions (White 1979; Ghigna et al. 1994), which can be expressed as sum of N-point correlation functions. Once the galaxy distribution has been smoothed with an appropriate filter, more topological features can be extracted from the isodensity surfaces. The results can be compared with theoretical distributions, e.g. Gaussian density fields, and non-linear properties can be detected. The topological genus statistic (Gott, Dickinson & Melott 1986) measures the relative abundance of overdense clusters to void regions and was estimated for the 2dFGRS (James et al. 2009) and a SDSS galaxy catalog (Gott et al. 2008), where on some scales slight deviation from the pure Gaussian field were found. An extension of this analysis, the Minkowski Functionals (Mecke, Buchert & Wagner 1994) have been calculated for galaxy surveys (e.g. Kerscher et al. 1998; Schmalzing & Diaferio 2000) like SDSS (Hikage et al. 2003) and galaxy cluster catalogs (Kerscher et al. 1997, 2001), where significant non-Gaussian features in the large-scale spatial distribution of galaxy clusters were found as well.

A particular technique to quantify the filamentary character of the galaxy clustering is based on the minimal spanning tree (MST) that was introduced by Kruskal (1956). The MST is a graph consisting of a set of N points (here galaxies) and the unique network of $N-1$ edges that link two points. If there is a path between any pair of edges, the graph is called connected and a connected graph with no circuits is called a tree. The MST is then the tree of minimal length that contains all points. In cosmology, MST was used to determine the statistical significance of filament-like features. This approach was used as a statistic to find clusters and filaments in the CfA redshift survey (Barrow, Bhavsar & Sonoda 1985) and to demonstrate that filaments are real and not just random alignments of galaxies (Bhavsar & Ling 1988).

1.5 Geometric Properties

Instead of quantifying the large-scale galaxy distribution with a global structure property, a further approach can be to consider local statistics of the galaxy distribution. This structural information

about the individual neighborhoods around galaxies can be related to galaxy properties that have been altered due to galaxy interactions. For instance, the estimation of the local number densities of galaxies within a specified scale or the number of nearest neighbors led to the well-known morphology-density relation (Hubble & Humason 1931; Dressler 1980). Local structure statistics can address questions like what structural types of environments occur and whether the local density alone is sufficient to distinguish between these differently shaped environments of galaxies. If all possible environments cannot be separated with the local density, then different structural features extracted from the local environment should be considered to provide the missing information. One can think of different local spatial distributions of galaxy locations that have different geometry features but the identical local density. Even for these environments, the galaxy s physical properties can still depend on the geometry of the local distribution, e.g. whether the galaxies is located in a filament-like or wall-like environment.

The existence of different geometries like clusters, filaments and wall-like structures, is motivated and predicted by theoretical implications, e.g. the Zeldovich approximation (Zeldovich 1970), numerical simulations (White et al. 1987; Springel et al. 2005), and was observed in the galaxy distribution (Doroshkevich et al. 2004). So far, the mentioned higher-order statistics (see Section 1.4) have been used to extend the description of these structures with density estimators. However, these global measures characterize the whole galaxy distribution due to averaging over the considered volume. It is not possible to assign each galaxies to a geometry class. Instead, recent approaches try to find regions in the galaxy distribution, which consist of specific geometries like filaments (Colberg et al. 2005).

To provide such a local description for galaxy surveys, an estimation of local geometry can be determined using the scaling index method (SIM, Jamitzky et al. 2001) that can provide a local scaling index α for each galaxy. This scaling index measures the local scaling property of neighbor galaxies within a given distance around the center galaxy. The scaling property distinguishes between galaxy neighborhoods of point-like, line-like, plane-like, and low-density geometry, which correspond to cluster, filament, sheet, and void-like elements of the large-scale galaxy distribution, respectively. The geometry can be characterized by a scale-dependent slope $\alpha(r)$ of the *local* galaxy-galaxy correlation function, $\xi(r) \sim r^{-[3-\alpha(r)]}$. For the estimation of $\alpha(r)$ a robust method was used with simulated dark-matter distributions (Raeth et al. 2002; Huber 2002) and cosmic microwave background data (Raeth & Schuecker 2003; Rath et al. 2007; Rossmanith et al. 2009).

A corresponding estimate to a *global* $\alpha(r)$ relation was first obtained from a two-point correlation function $\xi(r)$ in the strong nonlinear regime $r < 5\,h^{-1}\,\mathrm{Mpc}$ (Peebles 1980; Davis & Peebles 1983), where the slope of $\xi \quad r^{-\gamma}$, $\gamma = 1\ 8$, leads to $\alpha = 3-\gamma = 1\ 2$ independent of scale (e.g. Weinberg et al. 2004). However, global statistics average over all local relations of each galaxy and, thus, provide only information about the averaged and global clustering properties. Instead of this average, the local scaling behavior of each galaxy neighborhood can provide a more detailed information, which is now specific for each galaxy. This geometrical feature can be useful for the investigation of galaxy evolution, which can depend on the galaxy s environment. Since correlations between galaxy properties and environment extend to low-density regions and the outskirt of clusters, the question arises if the different geometric properties found in these regions are related to the galaxy properties as well.

The local geometry can support the understanding of the origin of the observed relations between galaxy properties and its environment. The standard cosmological paradigm is based on the model of hierarchical clustering (White & Frenk 1991), where the large-scale structure of the present-day Universe has evolved through merging from smaller structures to progressively larger ones. Density fluctuations on galaxy scales collapsed earlier and formed the over-dense regions in the Universe seen today. Galaxies in high-density regions such as galaxy clusters formed even earlier and are more evolved compared to galaxies in low-density regions such as void-like regions in space. Unlike these members of void-like regions, galaxies in dense regions are more likely to be influenced by processes and interactions like stripping or harassment, which are typical for these environments. By assigning a local geometry estimated by α to each galaxy, relations between observed galaxy properties and different classes of local environments can be compared. This approach provides

further tests about the origin of these relations, for instance to address the question whether these relations were established during the formation of galaxies (nature) or are the product of various processes during the galaxy s history (nurture).

1.6 Galaxy Surveys

In order to construct a three-dimensional map of the Universe, galaxy surveys measure the redshift of an object to measure its distance. The use of modern multi-fiber spectrographs increases the number of available redshifts rapidly and allows to create wide-angle redshift surveys with a large sky coverage. The Sloan Digital Sky Survey (SDSS, York et al. 2000) is the largest galaxy redshift catalog ever compiled and offers the opportunity to accurately measure the dependency of galaxy properties as a function of the environment. SDSS provides five-passband photometry with medium resolution spectroscopy for each galaxy. The used flux-limited sample from the fourth data release (Adelman-McCarthy et al. 2006) included roughly $3\,8 \times 10^5$ from a total of $4\,8 \times 10^5$ galaxies with measured redshifts at a median redshift of $z_{0\,5} = 0\,10$. The sky coverage of the spectroscopic data set reached $\sim$4800 deg^2. This large sample size and volume allow stable estimations of statistical measures of the local galaxy environments. In addition, the SDSS catalog includes many examples of the typical constituents of the large-scale structure, like cluster regions that are connected with filament-like structures. One of the previous large galaxy surveys, the 2dF galaxy redshift survey (2dFGRS) (Colless et al. 2003; Cole et al. 2005), has a similar median redshift $z_{0\,5} = 0\,11$, but neither the large number of observed galaxies ($2\,2 \times 10^5$) nor the same sky coverage ($\sim$1800 deg^2). The 6dF galaxy redshift survey (6dFGRS) (Jones et al. 2009) has a larger sky coverage of $\sim$17000 deg^2, but a lower median redshift $z_{0\,5} = 0\,053$ and less galaxies ($1\,6 \times 10^5$).

1.7 Outline

This work is organized as follows: Chapter 2 provides more details about the used SDSS data set along with the estimation of the photometric and spectroscopic galaxy properties. The local geometry estimation is presented in Chapter 3, where the structural meaning, possible distortion, and sources of errors are discussed. Chapter 4 presents the results including the relations between the galaxy properties and the local density and geometry. These result are compared and discussed in Chapter 5.

Chapter 2

SDSS Data

We describe the technical details of the SDSS data including the details about the used galaxy subsamples, the photometric and spectroscopic measurements, and the calculated structure parameters.

2.1 SDSS Details

The Sloan Digital Sky Survey (SDSS)[1] provides photometric and spectroscopic data covering over π steradians of the northern Galactic cap. The wide-field 2.5 m telescope at Apache Point Observatory, Sunspot, New Mexico, images the sky in five bands ($u\ g\ r\ i\ z$) between 3000 and 10000 Å (Fukugita et al. 1996; Gunn et al. 1998; Hogg et al. 2001; Smith et al. 2000). The used drift-scanning, large-format mosaic CCD camera (Gunn et al. 1998) detects objects to a flux limit of $r \sim 22\ 5$ mag.

One objective of SDSS is to obtain the spectra of almost 700 000 galaxies with a magnitude limit of $r = 17\ 77$ (Strauss et al. 2002). In addition, the spectra of 10^5 luminous red galaxies (Eisenstein et al. 2001) and 10^5 quasars at redshifts $z \lesssim 5\ 8$ selected for optical spectroscopy Richards et al. (2002) will be assessed. Many details about the spectroscopic target selection of the main galaxy sample are found in Strauss et al. (2002). The descriptions of the observation hardware, the data processing software, the measured quantities for each observed object, and an overview of the properties of this data set are found in Early Data Release paper (Stoughton et al. 2002).

The SDSS images are reduced and catalogs are produced by the SDSS pipeline PHOTO, which measures the sky background and the seeing conditions. It detects stellar objects and measures their photometric properties. The magnitudes are calibrated to a standard star network (Smith et al. 2000), which is close to the AB system (Fukugita et al. 1996). All the magnitudes in SDSS are asinh magnitudes (Lupton, Gunn, & Szalay 1999). Since the difference between the asinh magnitudes and conventional logarithmic magnitudes is negligible in our sample, we apply no conversion and treat asinh magnitudes as standard AB magnitudes.

For galaxy photometry, measuring flux is difficult because galaxies possess different radial surface brightness profiles and only poorly defined edges. The properties of the azimuthally averaged radial profile $I(r)$ of a galaxy can be expressed in a general form using Sérsic law, first introduced by Sérsic (1968):

$$I(r) = A \exp\left[-\left(\frac{r}{r_0}\right)^{1\ n}\right] \tag{2.1}$$

where A, r_0, and n are free parameters quantifying the amplitude, size, and shape of the surface brightness profile quantitatively. For $n = 4$, Equation (2.1) becomes de Vaucouleurs law and

[1] `www.sdss.org`

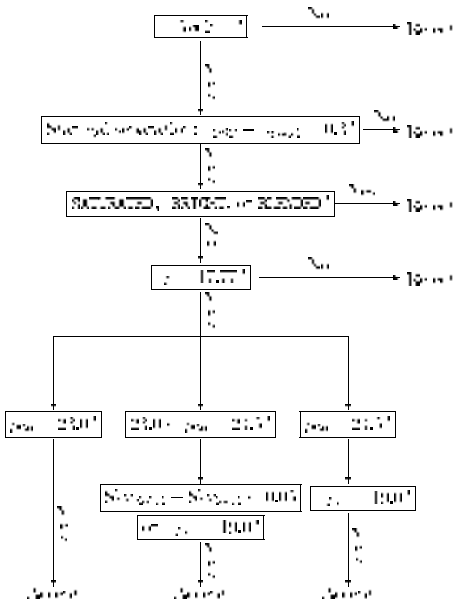

Figure 2.1: Schematic flow chart of the main galaxy target selection algorithm from Strauss et al. (2002).

characterizes the profiles of elliptical galaxies and bulge components of disk galaxies. For $n = 1$, Equation (2.1) is a exponential law describing the radial profiles of disk components of disk galaxies. As found by Blanton et al. (2003b), galaxies with exponential profiles tend to be blue with a low luminosity and low surface brightness, while quite concentrated galaxies in the range $n = 4 - 5$ tend to be red with a high luminosity and high surface brightness.

In order to avoid photometric biases, galaxy fluxes should be defined as a constant fraction of the total galaxy s flux independent from the position and the distance of the galaxy. Two object fluxes that fulfill this requirement are calculated by PHOTO, as described by Stoughton et al. (2002). The first flux is the SDSS Petrosian magnitude system, a modified version of the original quantity introduced by Petrosian (1976). The essential feature of these magnitudes is that in the absence of seeing a constant fraction of a galaxy s light is measured regardless of distance or size (see Blanton et al. 2001 and Strauss et al. 2002 for more details). The second appropriate measure of the flux of galaxies is the SDSS model magnitude, which is the best-fit model of a pure de Vaucouleurs and a pure exponential profile to the two-dimensional image of each object (Blanton et al. 2003b).

The spectroscopic target selection is described in Strauss et al. (2002). The SDSS fibers are allocated to three main samples of galaxies, luminous red galaxies and quasars. The whole schematic flow of the main galaxy target selection algorithm is given by Strauss et al. (2002) and is shown in Figure 2.1. The three major selection criteria are

$$r_{\rm PSF} - r_{\rm model} > s_{\rm limit} \quad r_{\rm petro} < r_{\rm limit} \quad \mu_{50} < \mu_{50\ \rm limit} \tag{2.2}$$

where $r_{\rm PSF}$ is the r-band magnitude obtained fitting a point-spread function to the object, $r_{\rm model}$ is model magnitude and $r_{\rm petro}$ the SDSS Petrosian magnitude and μ_{50} is the half-light surface brightness, which is defined as the average surface brightness within the radius that contains half of the Petrosian flux. The thresholds of the values vary across the sky in a well-understood way, but for the bulk of the area, the limits are $s_{limit} = 0\ 3$, $r_{limit} = 17\ 77$ and $\mu_{limit} = 24\ 5$ mag. All these quantities are corrected for galactic extinction according to the maps of Schlegel, Finkbeiner & Davis (1998). The spectra cover the rest-frame wavelength from 3800-9200 Å at median redshift with a resolution of $R \sim 2000$ and are obtained through a 3 diameter fibers. The k-corrections

	galaxies
all	382772
volume limited	93873
$(0\,05 < z < 0\,1)$	
r= 2 h^{-1}Mpc	62732
r= 3 h^{-1}Mpc	50647
r= 5 h^{-1}Mpc	34565
r= 8 h^{-1}Mpc	18575
r= 10 h^{-1}Mpc	11127
r= 15 h^{-1}Mpc	2283

Table 2.1: The number of galaxies included in volume limited subsamples for different radii r. Larger radii lead to smaller galaxy samples.

eliminate the redshift dependence of the mean galaxy colors (Hogg et al. 2002) and are calculated with the the routines given by Blanton et al. (2003a).

2.2 Volume Limited Subsample

For our analysis of the geometry of the local galaxy distribution, we defined a volume limited galaxy samples. If the usually larger flux limited sample was used, the selection function and local densities would become redshift dependent, which complicates the further analysis and the interpretation of the results. Therefore, we extracted a volume-limited subsample of the fourth data release (DR4) of the SDSS galaxies (Adelman-McCarthy et al. 2006).

To access the data, we used the SDSS Query Analyzer[2] (sdssQA), a graphical user interface that allows to compose SQL[3] database queries. We used the query given in Appendix B.2 to extract coordinates (redshift, right ascension, declination), photometric (model and petrosian magnitudes), and spectroscopic (Hα, Hβ, and [OII], see Section 2.7) properties of the galaxies in our sample. The chosen flags include only objects from the main galaxy sample, and those objects that inhere a galaxy-like spectroscopic signature.

The flux limited sample contains 382,772 galaxies that satisfy the requirement of our query. An equatorial slice of this subsample for galaxies with $z < 0\,11$ is illustrated in the upper panel of Figure 2.2. By recreating the volume limited sample (Figure 2.2, lower panel) including only galaxies with redshifts $0\,05 < z < 0\,11$ and absolute magnitudes $-24\,5 < M_r < -19\,9 + 5\log(h)$, 93879 galaxies remain in our catalog that are distributed in a volume of $1\,58 \times 10^7\,h^{-3}\,\mathrm{Mpc}^3$, which results in a mean number density of $n = 0\,0059\,h^3\,\mathrm{Mpc}^{-3}$. The sample depth is about $320\,h^{-1}$ Mpc ($z = 0\,05$) and ranges over $503\,h^{-1}$ Mpc ($z = 0\,11$) along the right ascension axis.

In addition to the mentioned selection criteria, we excluded galaxies close to survey boundary from the calculation of point statistics in the subsequent analysis. These galaxies introduce edge effects and bias the sample means especially for the estimation of the local geometry (see Section 3.1.3 for more details). To avoid this effect, we calculated the distances between each galaxy and the sample boundary, and labeled all galaxies with a distance greater than $2r$, where $r\ (h^{-1}\,\mathrm{Mpc}) = 2\ 3\ 5\ 8\ 10\ 15$ is a set of scales. In Table 2.1, we summarized the numbers of galaxies for the different samples with different inclusion criteria. In the following sections, local densities and geometry will be calculated for these various scales and galaxies close to the sample boundary can be identified with those labels.

[2] `http://cas.sdss.org/astro/en/help/download/sdssQA/`
[3] Structured Query Language

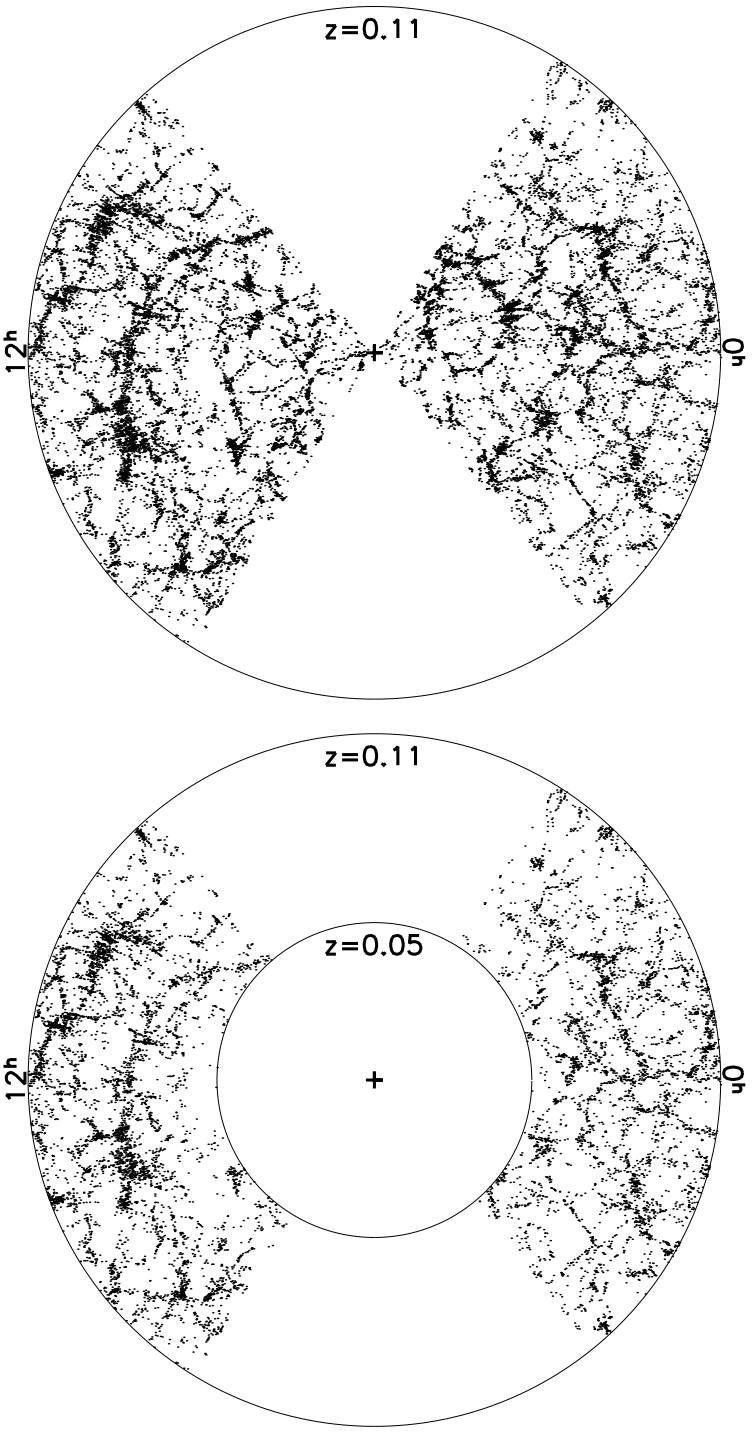

Figure 2.2: Two dimensional projections of equatorial slices for the flux-limited sample (upper panel) and volume limited sample (lower panel).

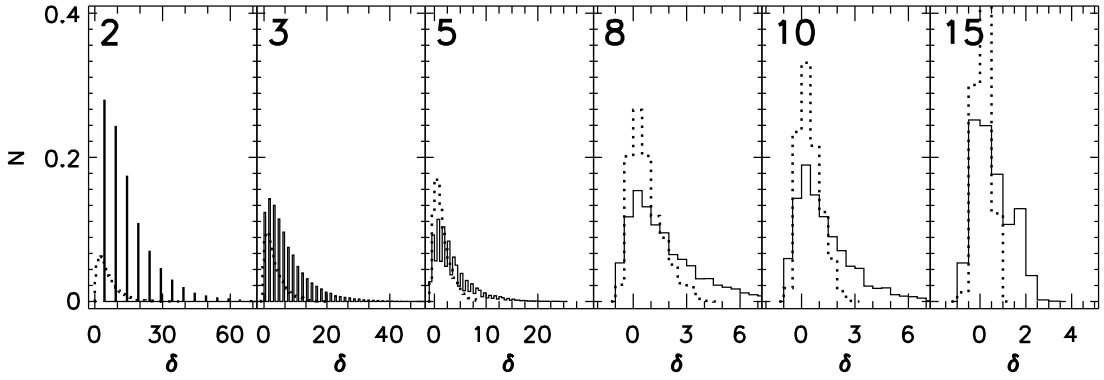

Figure 2.3: The density contrast with top-hat filter $\delta_{\rm th}$ (solid line) is plotted against the density contrast with Gaussian filter $\delta_{\rm g}$ (pointed line) for different radii (upper left corner of each subpanel).

2.3 Density Estimators

In this work, the local galaxy density is used as a reference parameter for the evaluation of the performance of the local geometry in its ability to characterize galaxy environments. Galaxies and the underlying matter density field are not homogeneously distributed in the Universe and, instead, show spatial fluctuations. In cosmological models, the global, volume averaged mean matter density ρ defines the dynamics and the fate the Universe, while local deviations from ρ and their evolution with cosmic time are fundamental predictors of the statistics of the large-scale galaxy distribution and galaxy formation (Peebles 1980; Peacock 1999). In cosmology, local variations from ρ at the position x are expressed through the density contrast $\delta^m(x)$ defined by

$$\delta^m(x) = \frac{\rho(x)}{\rho} - 1 \tag{2.3}$$

Here, we are mainly interested in the point statistics of the discrete galaxy distribution rather than the continous matter density field. The corresponding term for the number density contrast $\delta^g_r(x)$ of an arbitrary galaxy located at x can be written as

$$\delta^g(x\ r) = \frac{\rho(x\ r)}{n} - 1 \tag{2.4}$$

where the radius r is a free parameter and defines the size of sampling sphere that includes all neighbor galaxies that contribute to the local density. In general, a linear and deterministic relation between the density fluctuation fields of galaxies and matter is assumed, $\delta^g = b\,\delta^m$, where b is the galaxy biasing. However, more elaborate and nonlinear biasing schemes considering the physics of galaxy evolution are necessary to explain e.g. certain aspects of galaxy formation in high-density environments (Kaiser 1984; Mo & White 1996; Dekel & Lahav 1999; Taruya & Suto 2000; Seljak et al. 2005). These schemes indicate that higher order statistics are needed to fully describe the resulting structures in the galaxy distributions.

The local galaxy density gains further physical relevance in measuring the local galaxy interaction probability. The well-known morphology-density relation (Dressler 1980; Dressler et al. 1997; Goto et al. 2003b) relates the evolution of galaxies with the surrounding galaxy number density. Galaxy interaction occur more likely in high density regions like clusters, where galaxies are concentrated and strongly affected by the surrounding conditions. The estimation of the local density provides a map of the interaction intensities, where different density levels can be associated with different physical processes in galaxy evolution.

Different approaches have been used to estimate the local galaxy density which can lead to different physical interpretations. While the global density measures baryonic matter and luminosity content in the Universe, the interpretation of local densities estimation depends on the used weighting functions. Three major weighting functions can be found in astrophysical literature, namely the top-hat density, Gaussian-filtered density and adaptive density measures.

The standard top-hat density gives equal weight to galaxies within a certain radius r of the central galaxy for which the local density is measured. The advantage of this approach is that the physical scale, on which the density is measured, is identical with the radius and the physical interpretation is straightforward. However, the density of cluster regions are underestimated on scales larger than the typical cluster size ($\sim 5h^{-1}$Mpc). On very small scales, discrete number densities lead to a overestimation of the local density (e.g. Hogg et al. 2003).

Instead of using a top-hat filter, Gaussian densities weight neighboring galaxies with a Gaussian kernel. Close neighbor galaxies receive more weight and contribute more to the local density. The density field is smoothed and continuous, especially on small scales with small number statistic. In order to interpret the physical meaning, the kernel shape and the fuzzy edge of the filter volume must be considered. Gaussian kernels were used e.g. to calculate topological properties (genus statistics) of the large-scale structure (Gott, Dickinson & Melott 1986) and small-scale clustering (Hogg et al. 2003).

The third approach are adaptive density measures that estimate the local density on a scale dependent of the environment of the galaxy. For each galaxy, the n-nearest neighbor galaxies are identified, where typical values in cosmology range from $n = 3$ to $n = 10$. The distance of this nth-nearest neighbor determines the volume of interest and thus the local density. This method is adaptive to the density field in the sense that the parameters of the density parameters are not fixed (like the radius in the above two estimators) but are a function of the environment. This has the advantage that densities of clusters are given more accurately and are not underestimated like in the top-hat case. Adaptive measures are less noisier in low-density areas because the larger sampling volume smoothes the surrounding density field. The disadvantage is that one looses the physical scale dependence, an information crucial for environmental features. The comparison between surroundings of an ensemble of galaxies should be performed with caution since densities are measured with different volumes dependent on one galaxy.

Note that, although top-hat and Gaussian densities estimations have different properties, they have in common that they assume isotropy for the surrounding galaxy distribution. Both estimation take no account for the geometry of the galaxy distribution within the regions of interest. In contrast, the n-nearest neighbor density accounts for spatial anisotropies in the surrounding galaxy distribution. For example, these anisotropic environments are found in galaxies that are embedded in filaments. For those filament galaxies, the nearest neighbor density is more reliable density estimation. A comprehensive overview in the estimation of densities is provided by Silverman (1986).

In the cosmological context, it is interesting that galaxy properties depend on the local density. Neighboring galaxies influence the merging history and thus the evolution. The galaxy formation in galaxy clusters is different from galaxies formed in under-dense regions. Moreover, not only the observation, but the modeling of this environment dependent galaxy properties with simulations is an active field of research (e.g. Springel et al. 2005). In the following sections, a set of galaxy properties is presented, for which the dependence on the environment is investigated in the later analysis.

2.4 Luminosity

The apparent brightness of galaxies can be measured directly and is the fundamental property of any astrophysical object. The SDSS database provides measurements in five wave bands: $u\ r\ g\ i\ z$. We extracted the corresponding apparent magnitudes $m_u\ m_r\ m_g\ m_i\ m_z$ from the database and calculated the corresponding absolute magnitudes M. Following Peebles (1993), the

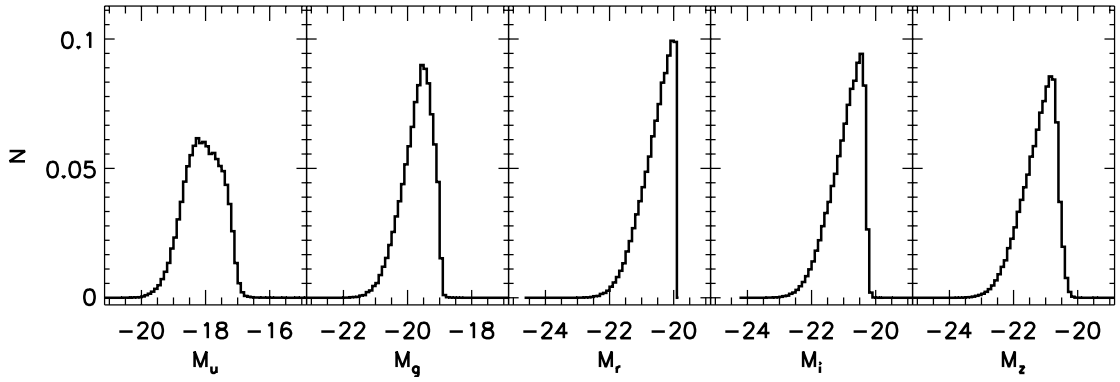

Figure 2.4: Histograms of absolute magnitudes for the five bands. M_u has the largest dynamic range and M_r has a sharp lower cut-off due to the volume limited sample.

comoving distance D_C is given by

$$D_C = D_H \int_0^z \frac{dz}{E(z\,)} \tag{2.5}$$

with the Hubble distance $D_H = cH_0^{-1}$ and the cosmological model dependent function $E(z) \equiv \sqrt{\Omega_M(1+z)^3 + \Omega_\Lambda)}$, with the matter density $\Omega_M = 0\ 3$ and the vacuum energy density $\Omega_\Lambda = 0\ 7$. The distance modulus DM is defined by the magnitude difference of an object s bolometric flux and it s apparent flux at 10pc:

$$DM \equiv 5 \log \left(\frac{D_C}{10pc} \right) \tag{2.6}$$

The absolute magnitude M is defined to be identical to an object s apparent magnitude m of 10pc distance:

$$M = m - DM - K(\lambda) \tag{2.7}$$

where $K(\lambda)$ is the k-correction (Oke & Sandage 1968; Blanton et al. 2003a). The k-correction adjusts the redshifted wavebands and is applied in order to compare the spectra of galaxies at different redshifts. In Figure 2.4, the five distribution functions for each absolute magnitude is shown. The wave-band M_u has the largest dynamic range and M_r has a sharp lower cut-off due to the volume limited sample.

2.5 Color

The color of a galaxy can be defined as the difference between two measurements of luminosity at different wavelengths. Galaxy colors quantify the stellar content of galaxies and allow to study their evolution history even for faint galaxies (Strateva et al. 2001). Like apparent magnitude, colors can be directly measured. From the absolute magnitudes presented in the previous section, the following colors are used: $u-r$, $g-r$, $r-i$, $r-z$ and $i-z$. In Figure 2.5, the corresponding distribution functions for these color band are shown.

The colors $u-r$ and $g-r$ are commonly used in previous studies of the galaxy distribution (Strateva et al. 2001; Baldry et al. 2004; Balogh et al. 2004b). For instance, Strateva et al. (2001) showed that $u-r$ is the best color band in order to distinguish two main galaxy population, the blue galaxies ($u-r < 2\ 3$) and red galaxies ($u-r \geq 2\ 3$). Qualitatively, the strong bimodal shape of the distribution function of the $u-r$ in Figure 2.5 indicates the existence of two underlying galaxy population. In the distribution function of $g-r$, this trend is still apparent, although not

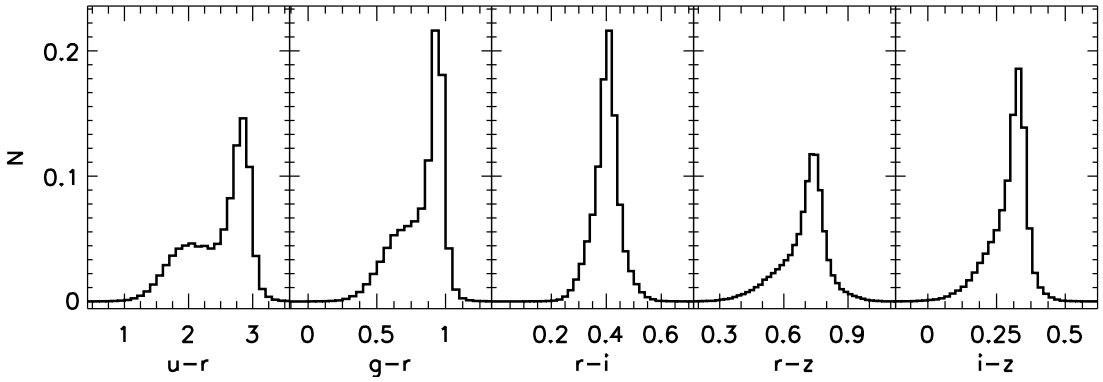

Figure 2.5: Colors histograms for different colors. Note that different bin sizes are applied.

that pronounced. The other three bands $r - i$, $r - z$ and $i - z$ were included to use the whole spectrum of wavebands range provided by SDSS.

The physical meaning of the colors for the galaxy evolution is qualitatively well understood and can be related to stellar populations. In this scenario, two main galaxy populations emerge in the Universe: blue and red galaxies. Blue galaxies are dominated by young, metal-rich stars with a lifetime of only $\sim 10^7$ years. In order to create new stars, galaxies need to cool and condensate gas clouds where the conditions for star formation are suitable. Note that, another reason for the identical galaxy color might be that all galaxies share the identical evolution history and are in the identical stage of their evolution. Because this is very unlikely, we can assume that these galaxies are actively star forming.

On the other hand, red galaxies are dominated by old red giants that are metal-poor. Thus, the star formation in those galaxies is not prominent or at a very low level such that the light of young stars can be neglected. There are various possible reasons for this lack of star formation, which seem to be very common in the Universe because these galaxies form a whole population. The most obvious could be the age of the galaxy: red galaxies are formed earlier than blue ones, thus their supply of gas is exhausted and no new stars can form. In this context, color is an indicator of age.

Besides this formation argument, there is the possibility of different evolution histories. Some galaxies may have formed in regions with less surrounding gas, e.g. where many galaxies in a small regions of space are supplied from the same gas reservoir. Other galaxies could have sufficient gas in there close proximity, but inhere not enough mass to drag the gas in their direction or do not provide the the conditions for condensations, which is a crucial requirement for star formation. Only cool and dense gas has a small enough Jeans length to clump together. This is a plausible scenario for galaxy clusters, where the hot intra-cluster gas and deep gravitational potentials omit gas cooling. In that case, the observed color reflects different galaxy evolution scenarios.

In the following sections we discuss star formation rate and morphology more closely. Yet already these qualitative arguments indicate that galaxy colors provide useful physical information about the dominant stellar population and the current star formation rate in the galaxy. They operate as indicators for the age of a galaxy and the abundances of metals and dusts. The differences in color of the two main galaxy populations indicate that galaxies are formed at different times, evolve differently or are located in different environments. The following parameters will reveil more accurate results.

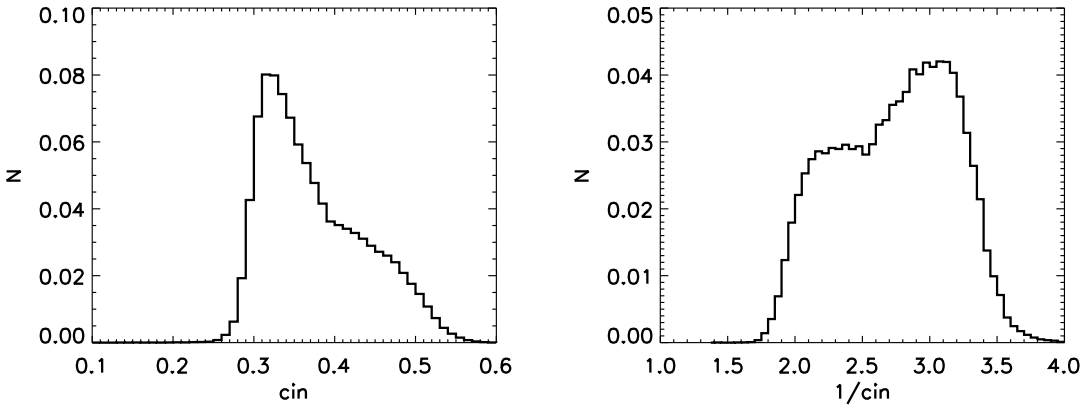

Figure 2.6: Histograms of the concentration index *cin* and its inverse 1 *cin*. Both definitions are related to the morphology of galaxies.

2.6 Morphology

Apart from the photometric analysis of stellar objects, the imaging of galaxies provides the possibility to classify the shape of the visible mass distribution, the morphology. Two main populations emerge in the Universe, which consist of two fundamental components of galaxy structure: a flat disk component of stars, gas and dust, and a spheroidal component, the bulge, which is a central, tightly packed group of stars that host one or more massive black holes in their center. Some galaxies include both constituents, but there are galaxies without a significant disk (ellipticals) or without a pronounced bulge (spirals and irregulars). The disk galaxies are subdivided into those with a high star formation rate (see next section) because of their rich abundance of gas and dust (spirals and irregulars), and those with no star forming activity (S0 galaxies). These observed morphologies of galaxies relate to different formation mechanisms and evolution models, e.g. the degree of interaction with the environment.

The different morphologies of galaxies can be classified by the well-known Hubble sequence (Hubble 1926), which mainly distinguishes between elliptical, spiral and irregular galaxies. These morphological types correlate with other galaxy properties, like the star formation rate (e.g. Kennicutt 1998a). The Hubble sequence is based on three characteristics: the bulge-to-disk-ratio, the tightness and the resolution of the spiral arms. There are two disadvantages involved with this method: first, it is hard to automate and therefore inefficient for large galaxy catalogues. Second, the Hubble sequence is not suitable for investigations of star formation because the resolution of the spiral arms depends particularly on the star formation rate. The structure of the spiral arms is more likely highlighted due to significant higher star formation activity along them.

Here, we use an automated method and choose an alternative classification, the concentration index *cin*, which is described e.g. by Morgan (1958), Shimasaku et al. (2001) and Goto et al. (2003a,b). This index is defined as the ratio of the radius containing 50% of the Petrosian flux to the radius containing 90% of the Petrosian flux as measured in the r-band:

$$cin = \frac{r_{50}}{r_{90}} \tag{2.8}$$

Low values of the concentration index ($cin < 0\ 3$) correspond to galaxies with a high central concentration of luminosity (small bulge-to-disk ratio), typical for early-type morphology. Spiral galaxies have a more uniform distributed luminosity because of the star-forming disk and are associated with higher values $cin > 0\ 4$. Shimasaku et al. (2001) analyzed properties of morphological classified SDSS galaxies and observed a strong correlation between the Hubble scheme and the

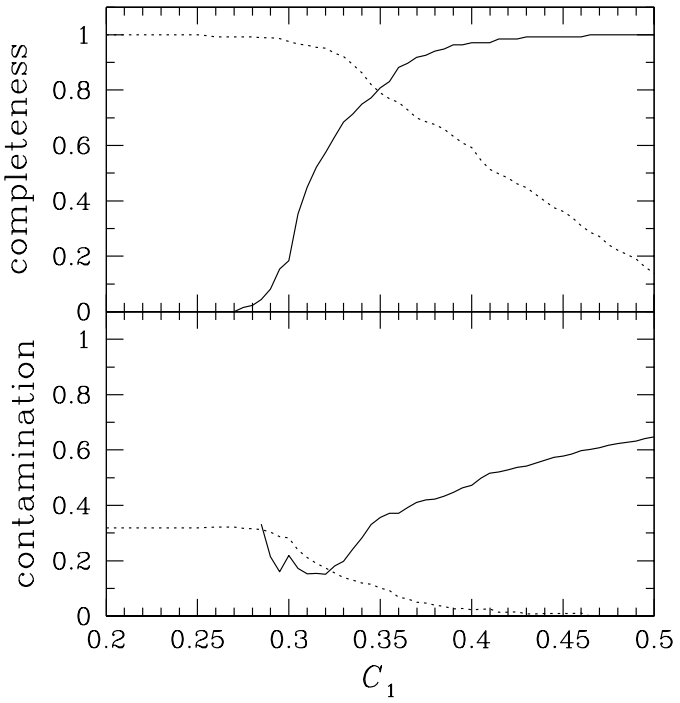

Figure 2.7: Completeness and contamination of early-type galaxies (solid line) and late-type galaxies (dotted line) classified by the concentration index C_1 (from Shimasaku et al. 2001).

concentration index *cin*. In Figure 2.6, the distribution function for *cin* for our galaxy sample is shown, as well the distribution function for the inverse concentration index 1 *cin*, which is used in the literature as well (e.g. Rojas et al. 2004). The bimodal shape of both histograms indicates the two main morphological types within the galaxy sample.

Using the concentration index, three different classes are defined to distinguish the morphological properties: all galaxies with $cin \leq 0\ 33$ are considered to be early-type, mixed type (irregular, S0) galaxies have a values in the range ($0\ 33 < cin \leq 0\ 375$) and late-type galaxies have the highest values ($0\ 375 < cin < 1$). Different choices for the threshold for *cin* lead to different completeness and contamination fractions of the morphology classes as shown in Figure 2.7 from Shimasaku et al. (2001). Our set of thresholds limits the contamination of the late-type galaxy sample to 5% at the expense of completeness. In general, the galaxy morphology correlates with the galaxy color. Spiral galaxies appear bluer than elliptical and S0 galaxies are red. This relation connects the physical processes on two different scales, the global shape of the galaxy at ranges of 100 kpc and the physics of star formation on scales of the solar systems. The few different shapes of galaxies show that there should be few major formation and evolution possibilities for galaxies in the universe.

Spiral galaxies have been able to collect gas from the surrounding reservoir and form disks with spiral arms. In these spiral arms, stars are formed, which end in supernovae in quite a short time (10^7 years). The shell from the stars is emitted and enriched the galaxy with metals. The large number of young stars in the pronounced disk leads to the blue light and the creation of metals. In contrast, elliptical galaxies possess no star forming disk. These galaxies were not able to form a disk or lost the disk during their evolution e.g. by a galaxy-galaxy interaction. This lack of young stars leads to a red color and a smaller metal abundance.

The dependence of these morphologies on the galaxy s environment is well-known (Hubble & Humason 1931; Dressler 1980; Postman & Geller 1984; Dressler et al. 1997). Specifically, Postman

& Geller (1984) observed that low-density fields ($\lesssim$ 5 galaxies Mpc^{-3}; $h = 1$) mainly consist of spiral galaxies; at a density $\lesssim$ 600 galaxies Mpc^{-3} the fraction of S0 becomes greater than the fraction of spirals and for densities $\gtrsim$ 3000 galaxies Mpc^{-3} the elliptical fraction rises steeply. One possible interpretation is that galaxies share a common type at the beginning and traverse through different types due to subsequent evolution. For instance, during the process called ram pressure stripping , a spiral s gas is removed and the galaxy turns into a S0 type (Gunn & Gott 1972). In order to constraint different theories of galaxy formation and evolution, observations are needed which compare the relation between several galaxy properties and the environment. We follow this intention and include the morphology parameter to our analysis.

In addition to the environment dependence, observations could find relations between the morphology of a galaxy and its star evolutions. Supernovae Type Ib, Ic and II are found only in Spirals and Irregulars, where they occur only in spiral arms or close to HII-regions, typical star formation regions. Those objects are young, massive population I stars. However, Supernovae Type Ia, the fate of old, low-mass Population II stars, are common in elliptical galaxies. If a Supernova Type Ia is found in spiral galaxy, it occurred in the central bulge region but within the spirals arms.

2.7 Measuring the Star Formation Rate in Galaxies

Galaxies are dynamical systems which can be classified by their ongoing rate of star formation or their history of star formation activity. Current research in cosmology investigates the relations between the properties of the whole galaxy and the properties of the inhabitant stars. We included this parameter to our analysis in order to link the purely statistical information about the environment of a galaxy with its stellar content. This approach addresses the question whether the stellar population of a galaxy predicts its environment or if the surrounding of the galaxy constraints certain types of stars or limits star formation.

2.7.1 Indicators for Star Formation

The star formation rate (SFR) of galaxies can be indirectly determined from observational data using various approaches (Kennicutt 1998a; Sullivan et al. 2001; Charlot & Longhetti 2001). The most accurate approaches are to utilize the correlation of SFR with measurements of the (1) far-infrared luminosity, which arises form dust-absorbed ultraviolet radiation from star formation, or (2) the 1.4 GHz radio luminosity which results from synchrotron radiation generated by relativistic electrons accelerated by Type II supernovae from stars of mass $\geq$7-8 $M_{\odot}$. Indicators that are sensitive to the ionized flux from massive stars ($\geq$5 $M_{\odot}$), which include measurements of the (3) ultraviolet continuum from young stars and the fluxes of (4) nebular emission lines, generated in regions ionized by the most massive early-type stars ($\geq$10 $M_{\odot}$). Each of these techniques suffers form different biases and calibration uncertainties but estimate the SFR of normal galaxies consistently if these effects are corrected (Hopkins et al. 2001).

In this work, we estimate the SFR based on the Hα recombination line (6564 Å), which is one the most dramatic differences in the integrated spectrum of galaxies of different types. All nebular lines effectively re-emit the integrated stellar luminosity of galaxies below the Lyman limit. Most analysis use the Hα line as a sensitive probe of the young massive stellar population (Lewis et al. 2002; Gómez et al. 2003), although other recombination lines, like e.g. Hβ, Pα, Pβ, are possible as well. The advantage of Hα emission lines are their high sensitivity and the direct coupling between the nebular emission and the SFR. However, the method is limited through the sensitivity to uncertainties in extinction and the initial mass function. In addition it assumes, that all star formation is traced by the ionized gas.

We also use the measurements of the forbidden emission line duplet [OII] λ3727 in order to compare the results with higher-redshift galaxies. The Hα emission line is redshifted out of the visible window beyond at $z \sim 0\,5$ and bluer, higher order emission lines are not reliable tracers of the SFR, because they are weak and influenced by stellar absorption. The [OII] line is the

strongest emission line in the blue regime of the spectrum. These excitations are not directly coupled to the ionizing luminosity. Furthermore, they are more sensitive to metallicity and the ionizing levels of the gas than the Hα emissions. However, empirical calibrations can provide quantitative SFR tracers, which is not done here.

In order to calculate the intensity of the spectral lines, we use their observed equivalent width (EW) listed in the SDSS database. The EW is the wavelength interval of the continuum spectrum, which includes the same total energy as the spectral emission line. In Figure 2.8, we show the histograms of the EWs.

The relation between SFR and ionized flux is usually computed using an evolutionary synthesis model, where only stars with masses of $>10\,M_\odot$ and lifetimes <20 Myr contribute significantly to the ionized flux. The emission lines provide an almost direct measurement of the SFR, independent of the previous star formation history. In order to calculate the SFR of the galaxies, we use the Hα fluxes calculated from line parameters measured by the SDSS spectroscopic data processing pipeline (Stoughton et al. 2002) and the theoretical model by Kennicutt (1998b):

$$\mathrm{SFR}_{\mathrm{H}\alpha}\,(M_\odot\,\mathrm{yr}^{-1}) = \frac{L(\mathrm{H}\alpha)}{1\ 27\times10^{34}\ W} \tag{2.9}$$

where $L(\mathrm{H}\alpha)$ is the observed luminosity in the Hα line. This relation is valid for the optical thin Case B recombination (no escape of Lyman α photons) with an initial mass function and a mass range from 0.1 to 100 $M_\odot$ (Salpeter 1955). This assumption may cause an underestimation of the SFR by a small factor if the line-emitting regions suffers from extinction (Charlot & Longhetti 2001). In addition, a burst-like star formation will effect the average SFR over short time scales ($\sim$ 100 Myr) (Sullivan et al. 2001). Neither of these effects will influence the following comparison of galaxy populations with similar luminosity functions. Kennicutt (1998b) provides a discussion about the effects of changing the mass scale or choosing a different mass function.

2.7.2 Corrections

Before applying the calibration in Equation 2.9, the Hα luminosities have to be corrected for stellar absorption and for obscuration due to intrinsic dust content of the target galaxy. In addition, aperture corrections have to be added, because the fiber diameter (3) can be smaller than the target diameter and emission from the galaxy can be lost. The corrections where applied following the prescription of Hopkins et al. (2003).

The Milky Way foreground obscuration is considered through the extinction corrections given by the dust maps from Schlegel, Finkbeiner & Davis (1998). Luminosity attenuation by the dust intrinsic to the star forming galaxy can lead to a more significant underestimation of the emission line flux. From the observed Hα and Hβ fluxes, the Balmer decrement $F_{\mathrm{H}\alpha}\ F_{\mathrm{H}\beta}$ is used to calculate the correction. The Balmer lines have to be corrected for stellar absorption, which induces an overestimation of implied obscuration (references in Hopkins et al. 2003).

The EW of Hα and Hβ are corrected for stellar absorption through adding the constant correction term $\mathrm{EW}_c = 1\ 3$ Å and $\mathrm{EW}_c = 1\ 65$ Å, respectively. This EW correction alters the the emission line flux:

$$S = \frac{\mathrm{EW} + \mathrm{EW}_c}{\mathrm{EW}}\,F \tag{2.10}$$

where S is the stellar absorption line flux for Hα or Hβ and F is the observed line flux. Although the median stellar absorption of the Hα line is 2.6 Å, the SDSS pipeline Gaussian fit underestimates the line emission by a factor of 2. Thus the correction of $\mathrm{EW}_c = 1\ 3$ Å seems appropriate. Finally, the dust attenuation was estimated through the flux ratio $S_{\mathrm{H}\alpha}\ S_{\mathrm{H}\beta}$, assuming that this intrinsic ratio is equal 2.86, and the Milky Way obscuration curve of Cardelli et al. (1989). Together, these corrections lead to a factor $(2\ 86^{-1}S_{\mathrm{H}\alpha}\ S_{\mathrm{H}\beta})^{2\ 114}$.

The aperture correction recovers an effective Hα luminosity for the whole galaxy based on the Hα EW together with an estimation of the continuum luminosity for a galaxy from the photometric

catalog at the observed Hα wavelength. The stellar absorption corrected Hα luminosities can be estimated by (Hopkins et al. 2003, Equation (5)):

$$L(\mathrm{H}\alpha)(W) = (\mathrm{EW}(\mathrm{H}\alpha) + \mathrm{EW}_c)10^{-0\ 4(M_r - 34\ 10)} \frac{3 \times 10^{18}}{[6564\ 61(1+z)]^2} \tag{2.11}$$

where M_r is the k-corrected absolute r-band AB magnitude which is calculated from the observed r-band Petrosian magnitude. The last term converts the luminosity units to W Å^{-1}. This correction assumes that fiber luminosity is representative for the whole galaxy and that the star formation is uniformly distributed within the galaxy. For the used volume-limited sample used in this work, the 3 aperture is equivalent to the median physical size of 4.1 kpc. The fiber probes a significant fraction of the galaxy, which is thus valid for the accuracy estimated by Hopkins et al. (2003). Compared with other SFR estimates from global galaxy properties, like radio, far-infrared or u-band continuum emission, this correction reproduces good results.

Including all corrections, the SFR indicated through the Hα emission line is given by (Hopkins et al. 2003, Equation (B2))

$$\mathrm{SFR}_{\mathrm{H}\alpha}\,(M_\odot\,\mathrm{yr}^{-1}) = \frac{L(\mathrm{H}\alpha)}{1\ 27 \times 10^{34} W} \left(\frac{S_{\mathrm{H}\alpha}\ \ S_{\mathrm{H}\beta}}{2\ 86}\right)^{2\ 114} \tag{2.12}$$

We used Equation (2.12) to estimate the star formation for each galaxy and investigated whether these results depend on the local environment of galaxies (Chapter 4).

2.7.3 Physical Motivation

Cosmic star formation varied during the evolution of the Universe. Early star formation was dominated by the creation of massive stars or star clusters because the gas cooling during the collapse was less effective compared to later stages. Hydrogen did not support the cooling process. Dust that contained heavier elements, which have higher cooling abilities, was not abundant yet. After the supernovae of the first stars, the interstellar medium was more fragmented and enriched with heavier elements. Both, the fragmentation of the depris and the heavier elements, increased the cooling efficiency of the matter, which could now form new, less massive stars.

Different types of star formation explain other properties of galaxies, e.g. the color. A high rate of star formation reflects the on-going formation of stars which emit mainly blue light, while red galaxies have a low rate of star formation. Although galaxy color is already an indirect measurement of star formation, the direct measurement of star formation provides us with a more accurate estimation on the amount of new stars and on the efficiency of star production in each galaxy.

The correlation between morphology and star formation is well-known: spiral galaxies are actively star forming, while elliptical galaxies ceased it some time ago. In order to investigate the mutual influence between environment and galaxy properties, it is still reasonable to include both properties. Both are driven by different physical processes that work on different scales. In addition, star formation is a more sensitive to transitions between galaxies with different stellar activity. The morphology of galaxies might not be sensitive for smooth transition between different stellar activity. While a galaxy s star formation can be affected by small changes that even occur in low-density regions, e.g. the supply of gas, the transformation of the morphology are most likely in high-density regions.

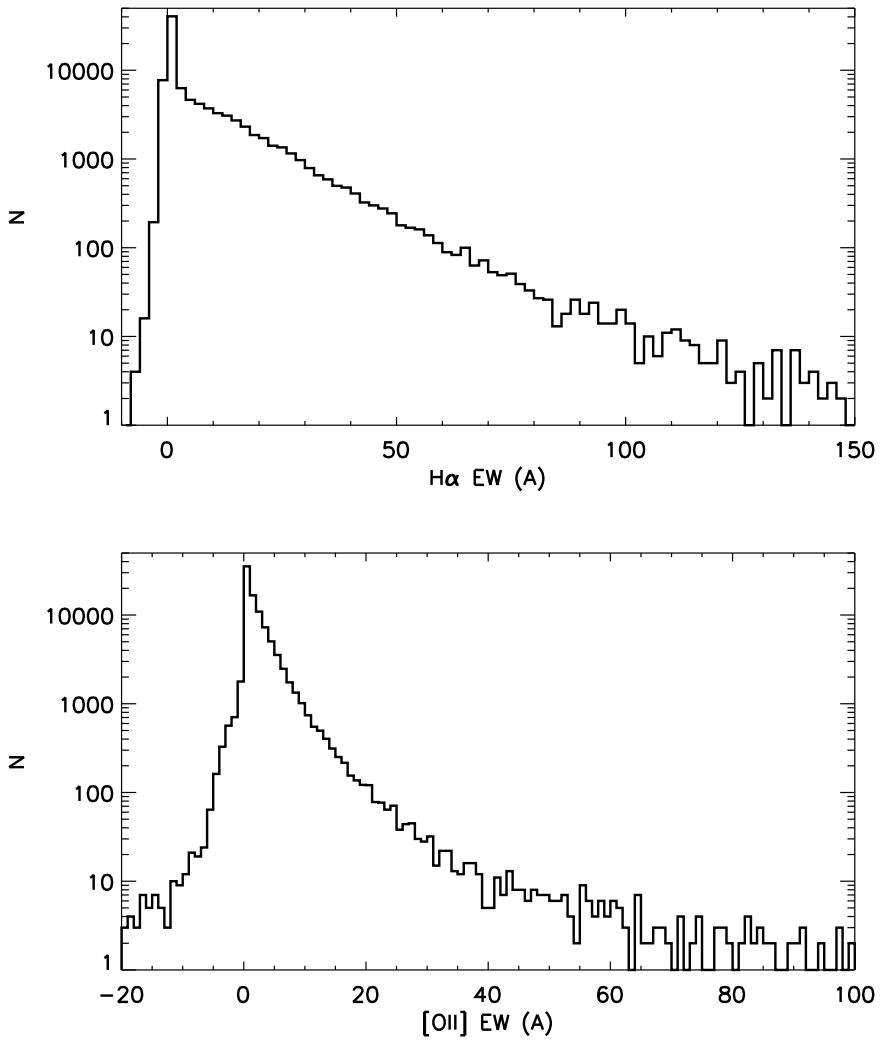

Figure 2.8: *Top*: Distribution of the observed Hα EW for the galaxies in our sample. *Bottom*: The distribution of the observed [OII] EW for the galaxies in our sample.

Chapter 3

Scaling Index Method

In this chapter, we introduce the geometry estimator α, which is derived form the two-point correlation function. We survey the possible values of α and test its stability. An alternative derivation of α can be found in the Appendix A.1.

3.1 Estimator of Local Geometry

3.1.1 Correlation Function

The definition of a stable estimator used for the classification of the geometry of the galaxy distribution around each galaxy is tightly connected to the well-known 2-point galaxy-galaxy correlation analyses. The classical form of this method, the angular 2-point galaxy-galaxy correlation function $w(\theta)$, was introduced by Totsuji & Kihara (1969) and Peebles (1973, 1980) to quantify the observed clustering of galaxies distribution projected on the celestial sphere. We can define $w(\theta)$ in such a way that the conditional probability $\mathrm{d}P_{12}$ of finding two galaxies in the solid angles $d\Omega_1$ and $d\Omega_2$ at a angular separation θ is given by

$$\mathrm{d}P_{12} = n_a^2 \left[1 + w(\theta)\right] \mathrm{d}\Omega_1 \mathrm{d}\Omega_2 \tag{3.1}$$

where n_a is the angular mean density of galaxies projected on the sky. For a 3-dimensional data set, an equivalent definition can be formulated for the spatial 2-point galaxy-galaxy correlation function $\xi(r)$:

$$\mathrm{d}P_{12} = n^2 \left[1 + \xi(r)\right] \mathrm{d}V_1 \mathrm{d}V_2 \tag{3.2}$$

where two galaxies are now found in the Volumes dV_1 and dV_1 separated by the distance r and n corresponds to the mean space density. To calculate the correlation function $\xi(r)$, the neighbor galaxies of all galaxies within the distance range $[r\ \ r{+}\mathrm{d}r]$ are counted and the results are averaged over the whole sample. In this way, $\xi(r)$ estimates the excess of clustering relative to a Poisson distribution, which can be created by a random point process for a given density. In addition, $\xi(r)$ provides a mean and global scaling behavior of the galaxy sample for distance range $[r\ \ r{+}\mathrm{d}r]$. This global scaling behavior for can be categorized into three cases:

- if $\xi(r) = 0$, the galaxies are uniformly distributed (Poisson distribution);
- if $\xi(r) > 0$, the galaxy locations are correlated and embedded in a cluster pattern;
- and if $-1 \leq \xi(r) < 0$, the galaxies are anti-correlated, which occurs on scales typical for the low-density regions in space (voids).

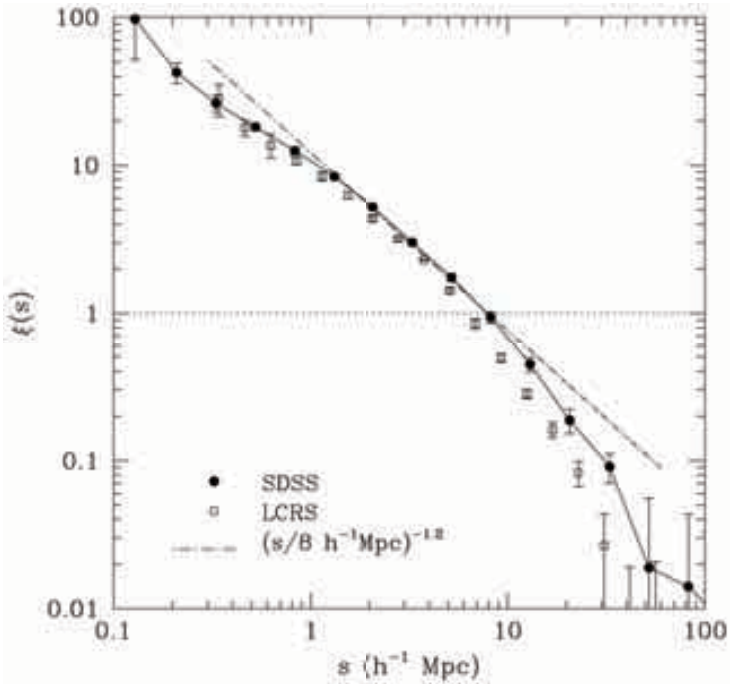

Figure 3.1: SDSS redshift-space galaxy correlation function $\xi(s)$ (filled circles, solid line) from Zehavi et al. (2002). A power-law fit for the range $2\,h^{-1}\mathrm{Mpc} < s < 10\,h^{-1}\mathrm{Mpc}$ is plotted as a dot-dashed line. Open squares show $\xi(s)$ obtained from the Las Campanas Redshift Survey (LCRS, Tucker et al. 1997).

It is convenient to compute the correlation function ξ as (e.g. Peebles 1980; Landy & Szalay 1993)

$$\xi(r) = \frac{N_{\mathrm{obs}}}{N_{\mathrm{rand}}} - 1, \tag{3.3}$$

where N_{obs} is the number of galaxy pairs in the observed sample with separations between r and $r + \Delta r$, and N_{rand} is the number of pairs in the same interval for the same number of galaxies distributed randomly over an identical area. Observational determination of ξ for large-scale galaxy distribution yields relationships of a power-law form:

$$\xi(r) = (r/r_0)^{\gamma}, \tag{3.4}$$

which is characterized by a typical length scale r_0 and the correlation dimension γ. Estimated values for these parameters are $r_0 = 5.4\,h^{-1}\mathrm{Mpc}$ and $\gamma = -1.77$ in the range $r \leq 10\,h^{-1}\mathrm{Mpc}$ for the CfA redshift survey Davis & Peebles (1983) and

$$\xi_{\mathrm{SDSS}}(r) = (r/r_0)^{-1.75},\ r_0 = 6.1\,h^{-1}\mathrm{Mpc} \tag{3.5}$$

in the range $0.1\,h^{-1}\mathrm{Mpc} < r < 16\,h^{-1}\mathrm{Mpc}$ for a SDSS galaxy sample (Figure 3.1, Zehavi et al. 2002, 2004). These results confirm the model of hierarchical galaxy formation and indicate the foam-like structure of the three-dimensional galaxy distribution.

Both parameters, r_0 and γ, characterize the galaxy clustering. The parameter r_0 marks the scale where ξ is equal to unity. For scales below r_0 ($\xi > 0$), the galaxy clustering must be considered to be in the nonlinear regime, whereas for scales above r_0 ($\xi < 0$), the clustering patterns can be modeled with linear theory. Thus, the typical length scale r_0 estimates an upper limit for typical structure size of the nonlinear objects, mainly the size of the virialised clusters. The exponent γ reflects the scaling properties and the self-similar nature of the galaxy distribution. The observed value $\gamma \sim 1.7$ suggest that galaxies cluster mainly located along one dimensional structures like

filaments ($\gamma = 1$) and two dimensional structures like walls ($\gamma = 2$). More details about the correlation function analyses, like different estimators and errors, can be found in e.g. Peebles (1980) and Martínez & Saar (2002).

The correlation function ξ is important for the theoretical description of cosmic structure formation. Linear analytic solutions for ξ can be calculated for the galaxy distribution at $z = 0$ from the Gaussian density fluctuations of the primordial Universe at $z \sim 1000$. The correlation function ξ, however, has some limitations in providing a complete description of the clustering properties of the galaxy distribution. The technique is a measurement of second order, i.e. an arbitrary point distribution is completely described by the correlation function, only if all moments above the second moments are zero. The point distribution is then completely described by its mean clustering and its variance of clustering. If ξ completely describes the clustering properties of galaxies, the galaxy distribution must be of second order and can be modeled with a Gaussian density distribution.

According to Equation 3.2, the calculation of ξ assumes that galaxies are isotropically distributed around an arbitrary galaxy. This assumption approximates well the clustering properties on global scales $\gtrsim 10\,h^{-1}$Mpc larger than the typical sizes of observed structures. On smaller scales, however, the gravitational force cannot be modeled by linear approximations which must be replaced by approaches that incorporate nonlinear terms, e.g. the Zeldovich approximation (Zeldovich 1970). This nonlinear evolution on small scales led to locally dense structures in the galaxy distribution like clusters and filaments, where galaxies are not isotropically distributed. Hence, the correlation function ξ cannot be considered to be a complete measurement of the local clustering properties. Some problems can be solved by studying higher order moments, but these can be difficult to compute.

Further limitations of the correlation function include the assumption that the galaxy distribution becomes spatially homogenous on a length scale that can be smaller than that of the catalog being analyzed. If this assumption is invalid, the amplitude of the correlation function and the length scale r_0 become scale dependent (Pietronero 1987). The correlation function is also limited by fact that it averages over amplitudes on a given scale that may come from galaxy pairs located different environments. This can yield a large amplitude for ξ because of a large number of close neighbors for a small fraction of the galaxies in the sample or because of smaller number of neighbors for nearly all galaxies in the sample. The correlation function does not indicate how many galaxies do contribute to the estimated value of clustering. Finally, this statistic is mostly applied either to all galaxies or to galaxies of the same properties like color or morphology (e.g. elliptical-elliptical pairs). Some environmental influences on galaxy properties might only be understood if all galaxies in an environment of a particular type are considered.

In the following sections, we develop a technique to describe the distribution of galaxies that originated in the field of nonlinear dynamics. Instead of averaging over the environment and clustering property of all galaxies and measure the excess of clustering relatively to a random distribution (Equation 3.2), we calculate the slope of the local correlation function of each galaxy. Different galaxies can possess different scaling relations because they are embedded in different environments. The two dimensional slice in Figure 2.2 shows that the prominent clusters, filaments and voids are part of the cosmic web. It is interesting to assess a local scaling behavior, i.e. a local scale length r_0 and a local exponent γ, which appear to be different for galaxies in different environments. This new local information can be crucial for the understanding of galaxy evolution and formation histories, which are both local processes that created different types of galaxies in different environments throughout the Universe. Although the correlation function does not provide local scaling information, it can be used to derive a geometry estimator that includes this local scaling behavior.

3.1.2 Global Geometry

In order to introduce a global scaling relation $\alpha(r)$ for the galaxy distribution, we use the 2-point correlation function $\xi(r)$ to calculate the mean local density n within a sphere of radius r around

a representative galaxy:

$$n(r) \;= \bar{n}\,[1+\xi(r)] \tag{3.6}$$

where $\cdot$ represents the sample mean and $\bar{n}$ is the mean density of the sample (see Chapter 2). If no clustering is present in the galaxy sample, the sample mean of the local densities $n(r)$ and mean density $\bar{n}$ are identical. To calculate the mean number of galaxies $N(r)$ contained in sphere with radius r around galaxies, $n(r)$ is integrated over the sample Volume V of the sphere:

$$N(r) = \int_0^r n(r\,)\;\mathrm{d}V(r\,) = \bar{n}\int_0^r [1+\xi(r\,)]\,\mathrm{d}V(r\,) \tag{3.7}$$

In the nonlinear regime with $\xi(r) > 1$ and $r < 30\,h^{-1}$Mpc, the galaxy correlation function $\xi(r)$ is very close to have a constant slope and can be approximated by a power-law of the correlation dimension γ,

$$1+\xi(r) \quad r^{-\gamma} \tag{3.8}$$

This self-similar behavior for galaxy clustering suggests that the concept of fractal geometry may apply (Mandelbrot 1982; Peebles 1993). Various fractal analyses have been applied to describe the self-similar behavior, such as wavelet transforms (e.g. Martinez et al. 1993) and percolation analysis (e.g. Klypin & Shandarin 1993). Here, we use the self-similar behavior to calculate the mean number of neighbor galaxies N for all galaxies in the sample. The integration over a sphere with the radius r then yields:

$$N(r) \quad \bar{n}\int_0^r r^{\,-\gamma}\,\mathrm{d}V(r\,) = \frac{2\pi\bar{n}}{3-\gamma}\,r^{3-\gamma} \tag{3.9}$$

Thus, the local number of galaxies follows a power law, $N(r) \quad r^{3-\gamma}$. In addition, a linear expression can be found for the relation between numbers of galaxies in two spheres around a representative galaxy, $N(r_1)$ and $N(r_2)$, with the given radii r_1 and r_2 ($r_2 > r_1 > 0$) by using logarithmic differences (e.g. Best et al. 1996):

$$\frac{N(r_2)}{N(r_1)} = \left(\frac{r_2}{r_1}\right)^{3-\gamma} \qquad \frac{\ln N(r_2) - \ln N(r_1)}{\ln r_2 - \ln r_1} = 3-\gamma \tag{3.10}$$

The correlation dimension is the slope of the functional relation between $\ln N(r)$ and $\ln r$ and describes how the number of neighbor galaxies scales with the radius. For the given radius interval $r_1 < r < r_2$, we call the scaling relation between the $\ln N(r)$ and $\ln r$ the global scaling index $\alpha(r)$,

$$\alpha(r_1\;\; r_2) = 3-\gamma \tag{3.11}$$

which is merely a different definition for a correlation dimension. Instead of calculating the logarithmic differences between a chosen radius interval, a more general expression of this global scaling index $\alpha(r)$ can be derived by performing a logarithmic derivation of Equation (3.9):

$$\alpha(r) = \frac{\mathrm{d}\ln N(r)}{\mathrm{d}\ln r} = \frac{r}{N(r)}\frac{\mathrm{d}N(r)}{\mathrm{d}r} = 3-\gamma \tag{3.12}$$

where the requirement $N(r) > 0$ holds for $r > 0$, since the galaxy in the center of the sampling volume is included in $N(r)$. For the given radius interval, the derived scaling relation is radius independent, $\alpha(r) = \alpha$. In general, the function $\alpha(r)$ values obtained at different scales r constitute to an overall scaling relation of the sample, $\alpha = \;\alpha(r)\;$, the mean scaling relation of all galaxies in the sample and all possible volume spheres defined by the radius r. This overall scaling index can be regarded as an galaxy distribution s fractal dimension that was derived similar to a box-counting approach. In Appendix A.1, more details about this approach are presented.

The observed global value of $\gamma = -1\;8$ corresponds to a scaling index of $\alpha(r < 30\,h^{-1}\mathrm{Mpc}) = 1\;2$, which is between the correlation dimension of filamentary structures ($\alpha = 1$) and wall-like structures ($\alpha = 2$). In the homogenous regime at scales $r > 30\,h^{-1}$Mpc , we have $\xi(r) \approx 0$ and

the correlation dimension becomes $\alpha(r > 30) = 3$, i.e. the dimension of the configuration space. Those examples indicate that the values of the scaling indices have an easy interpretation for the overall structural organization of the galaxy distribution. Since the relative positions of the galaxies are described by the scaling index $\alpha(r)$, which can be related to specific patterns, this estimated structural property of a point distribution can be called geometry.

However, a radius independent and global scaling relation might not adequately represent all existing structural information of the galaxy distribution. Based on the examples mentioned above, we expect to measure different scaling indices indicating different geometries at different radius intervals. The SDSS correlation function in Figure (3.1) shows that the correlation function should not be approximated over the whole range of radius r by one correlation dimension alone. Only certain radius intervals can be approximated with a specific correlation dimension and, according to our derivation, with a specific scaling index $\alpha(r)$.

$$\alpha(r) = 3 - \gamma(r) \tag{3.13}$$

With the mean value of the global scaling index, certain scaling regimes cannot be identified. Instead, a radius dependent scaling index $\alpha(r)$ classifies different scaling regimes and a spatial dependent scaling index could even classify galaxy environments with different scaling properties.

3.1.3 Local Geometry

The global scaling relation $\alpha(r)$ can be considered as a measure for the average scaling of all local scaling relations. Instead of calculating the sample property, these local scaling relations provide the possibility to characterize the local geometry *around each galaxy* in the sample. For the local case, we need to compute the local number of galaxies $N(\mathbf{x}_i\ r)$ for each of the N galaxies with comoving positions $\mathbf{x}_1 \quad \mathbf{x}_N$ and within a sample volume defined by the scaling radius r. Then, we can calculate the local scaling relation $\alpha(\mathbf{x}_i\ r)$. The corresponding relations between the global and local properties become

$$N(r) = N^{-1} \sum_{i=1}^{N} N(\mathbf{x}_i\ r) \quad \text{and } \alpha(r) = N^{-1} \sum_{i=1}^{N} \alpha(\mathbf{x}_i\ r) \tag{3.14}$$

To calculate the local properties $N(\mathbf{x}_i\ r)$ and $\alpha(\mathbf{x}_i\ r)$, we follow the derivation of the global scaling index in the previous section. However, we assume now that each local environment around each galaxy in the sample must be described with a local correlation function $\xi(\mathbf{x}_i\ r)$ and a local correlation dimension $\gamma(\mathbf{x}_i\ r)$:

$$1 + \xi(\mathbf{x}_i\ r) \quad r^{-\gamma(\mathbf{x}_i\ r)} \tag{3.15}$$

This approach is motivated by observations of large-scale structure of the Universe, where galaxies are located in different structural parts of the cosmic web. Galaxies in clusters of galaxies have a different local correlation dimension than galaxies located in filament-like shaped connections between the clusters. Like in the global case in Equation (3.12), we have to calculate the logarithmic derivative of the local numbers of galaxies $N(\mathbf{x}_i\ r)$ in order to assess the local geometry $\alpha(\mathbf{x}_i\ r)$ for each galaxy:

$$\alpha(\mathbf{x}_i\ r) = \frac{\ln N(\mathbf{x}_i\ r)}{\ln r} = \frac{r}{N(\mathbf{x}_i\ r)} \frac{}{r} N(\mathbf{x}_i\ r) \tag{3.16}$$

To solve this equation, we first need to define an analytically stable estimator for $N(\mathbf{x}_i\ r)$. For the global scaling index, we assumed that the average number of neighbor galaxies can be estimated with a power-law that holds in the nonlinear regime (Equation 3.9). If we calculate the local number of number of neighbor galaxies, this assumption might not be accurate for all environments; for instance, galaxies in low-density regions (voids) contain only a few or no neighbor galaxies with a given range. Rather than estimating a global average, we count the neighbors for each galaxy for a given volume to assess the slopes of the local correlation function and the local geometry. This counting leads to a discrete function $N(\mathbf{x}_i\ r)$ that is analytically unstable with respect to the

derivation operator r in Equation (3.16). Instead, the galaxy distribution can be represented by a smoothed density field that is continuously differentiable.

To derive such an analytic stable estimator for $N(\mathbf{x}_i\ r)$, we introduce a density function $n(\mathbf{x})$ to represent the discrete galaxy locations,

$$n(\mathbf{x}) = \sum_{j=1}^{N} \delta_D(\mathbf{x} - \mathbf{x}_j) \tag{3.17}$$

using Dirac s delta $\delta_D(\cdot)$, and a weighting function $w(d_{ij}\ r)$ to get the local number of galaxies $N(\mathbf{x}_i\ r)$,

$$N(\mathbf{x}_i\ r) = \int_0 \quad n(\mathbf{x})\, w(d_{ij}\ r)\, \mathrm{d}^3x = \sum_{j=1}^{N} w(d_{ij}\ r) \tag{3.18}$$

with an index $i = 1$ N and the Euclidean distance $d_{ij} = \mathbf{x}_i - \mathbf{x}_j$ between the ith and the jth galaxy. One possible choice for the weighting function $w(d_{ij}\ r)$ is a top-hat filter, which is defined by the Heavyside function $\Theta(\cdot)$:

$$w_{\mathrm{t}}(d_{ij}\ r) = \Theta(r - d_{ij}) = \begin{cases} 1 & : \text{for } 0 \leq d_{ij} < r \\ 0 & : \text{elsewhere} \end{cases} \tag{3.19}$$

This window function $w_{\mathrm{t}}(d_{ij}\ r)$ assigns equal weights to all galaxies within the distance r and was already used to calculate the global scaling index. In general, any window function and any distance measure could be used to calculate $N(\mathbf{x}_i\ r)$. To obtain analytic stable expressions, we use a set of Gaussian window functions, which assign neighbor galaxies different weights depending on the distance to the center galaxy:

$$w(d_{ij}\ r) = \exp[-(d_{ij}\ r)^q] \tag{3.20}$$

where the exponent q controls the weighting of the neighbor galaxies according to their distances to the galaxy for which α is calculated. Increasing q leads to a more stepwise weighting function and in the limit q , the weighting is equal to the weighting of the Heaviside function Θ. In the case $q = 2$, the window function is the well-known Gaussian exponential function. Throughout this work, we will use the exponent $q = 2$ to calculate α-values.

Inserting the expression for $w(d_{ij}\ r)$ (Equation 3.20) in the definition for $\alpha(\mathbf{x}_i\ r)$ (Equation 3.16) yields the following estimator for α:

$$\alpha(\mathbf{x}_i\ r) = \frac{2 \sum_{j=1}^{N} (d_{ij}\ r)^2 \exp[-(d_{ij}\ r)^2]}{\sum_{j=1}^{N} \exp[-(d_{ij}\ r)^2]} \tag{3.21}$$

This estimator for α was tested with simulated dark matter distributions (Huber 2002; Raeth et al. 2002), maps of cosmic microwave temperature fluctuations (Raeth & Schuecker 2003; Rath et al. 2007), and texture analysis (Huber et al. 2009). The following sections illustrate the meaning of α-values for specific geometry prototypes, such as filament-like and void-like regions, that typically occur in the galaxy distribution.

An advantage of the estimator in Equation (3.21) is the stability in low-density regions, i.e. regions containing no or only a few galaxies. Even if galaxies have no neighbors within a certain range, Equation (3.21) still provides analytically stable and reasonable results and assigns the galaxy a point-like environment ($\alpha \sim 0$). With this important feature, it is possible to identify galaxies located in void-like regions and compare them with galaxies in different environment types. With a relative abundance of $\sim 8\%$, void galaxies represent a group of galaxies with distinct physical properties and evolution histories compared e.g. to galaxies located in clusters (Rojas et al. 2004, 2005).

A limitation of this method is the sensitivity to edge effects especially for large radii. The survey boundaries mimic an artificial environmental geometry, which has no physical meaning.

For a given radius r, galaxies closer than $2r$ to those edges have to be remove in a conservative manner (minus estimator) like described in in Chapter 2. For the largest radii, the number of galaxies in the final sample used for the analysis shrinks down to ~ 2000. The large size of the SDSS galaxy sample allows the calculation of scaling relation for such large radii. Further must be noted, that in general, there is no analytical expression for the local geometry rather than a numerical result. The best choice for the radius parameter can depend on the data set and the objective, what kind of geometries are of interest. Another degree of freedom can be introduced by using an anisotropic distance measure.

3.1.4 Anisotropic Scaling Indices

The optimal selection of a filtering scale, e.g. to estimate the density of a point process, is a well-known problem in spatial statistics (e.g. Silverman 1986). For the large-scale galaxy distribution, the local density of a galaxies in clusters or filaments can be under-estimated if the scaling radius is larger than the diameters of the structures. In contrast, the local density of void galaxies can be over-estimated if the scaling radius is much smaller than the diameter of the void-like region. Since the best filtering scale can depend on the geometry of a local structure, a unique, optimal and overall scaling range for a galaxy sample does not hold the best classification results for all these structures.

Furthermore, instead of using one isotropic distance measure, which is applied for all galaxies, one can think of an adaptive intensity estimator. This estimator can be dependent on the position of the galaxy or dependent on certain properties of the galaxy. Here, we want to introduce such a measure in order to contribute to the problem of the redshift space distortions, which are discussed in the next section.

In general, the distance d_{ij} between two arbitrary galaxies located at the positions $\mathbf{x}_i$ and $\mathbf{x}_j$ can be written as

$$d_{ij}^2 = \mathbf{d}_{ij}^{\mathrm{T}}\, \mathbf{A}\, \mathbf{d}_{ij} \tag{3.22}$$

where $\mathbf{d}_{ij} = \mathbf{x}_i - \mathbf{x}_j$ is the distance vector between the two galaxies. The matrix $\mathbf{A}$ contains the distance metric; in the special chase of an unit matrix, $\mathbf{A} = \mathrm{diag}(1\ 1\ 1)$, the metric is constant and isotropic and the distance d_{ij} is equal to the Euclidean distance. All points with a given equal distance from a galaxy are found on the surface of a sphere around the galaxy. With a different choice of $\mathbf{A}$ with non-diagonal elements, the distance measure d_{ij} becomes non-isotropic and has an orientation. All points with a given equal distance from a galaxy are then found on a non-spherical surface.

Here, we want to define a local distance metric that depends on the location of the galaxy, the right ascension γ and the declination θ. Equidistant points are then located on the surface of a ellipsoid. We first define an anisotropic distance metric for the whole sample, where

$$\mathbf{\Lambda} = \mathrm{diag}(\lambda_1\ \ \lambda_2\ \ \lambda_3) \tag{3.23}$$

represents the metric with the eigenvalues λ_1, λ_2 and λ_3. If $\lambda_1 = \lambda_2 = \lambda_3$, the metric defines an isotropic distance measure, like used in Section 3.1.3. For all other cases with λ_1, λ_2, $\lambda_3 > 0$, the sampling volume is no longer a sphere but an ellipsoid. The semi-axis of this ellipsoid e_1, e_2 and e_3 are given by the eigenvalues of the metric $\mathbf{\Lambda}$:

$$\lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2 = const \tag{3.24}$$

where the relation between the lengths of the three semi-axes is

$$\frac{e_i}{e_j} = \sqrt{\frac{\lambda_j}{\lambda_i}} \quad i\ j = 1\ 2\ 3 \tag{3.25}$$

In the following analysis, we will use the following set of ellipsoidal distance metrics $\mathbf{\Lambda}_\lambda$ defined by one free parameter λ:

$$\mathbf{\Lambda}_\lambda = \quad \mathbf{\Lambda}\ \ \mathbf{\Lambda} = \mathrm{diag}(1\ \ \lambda^2\ \ \lambda^2) \tag{3.26}$$

where the relation between the long semi-axis and the two short semi-axes is $e_1 \;\; e_2 = e_1 \;\; e_3 = \lambda$. Then, the orientation of given distance metric $\mathbf{\Lambda}$ is rotated according to the right ascension γ and the declination θ:

$$\mathbf{A} = \mathbf{A}_\lambda(\gamma \;\; \theta) = \mathbf{R}^{\mathrm{T}} \, \mathbf{\Lambda}_\lambda \, \mathbf{R} \tag{3.27}$$

where the rotation matrix $\mathbf{R} = \mathbf{R}(\gamma \;\; \theta)$ includes the rotation around the y-axis ($\gamma = 0^\circ \;\; \theta = 0^\circ$) followed by a rotation around the z-axis ($\theta = 90^\circ$):

$$\begin{aligned}
\mathbf{R} &= \mathbf{R}_\phi \mathbf{R}_\theta \\
&= \begin{pmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix} \\
&= \begin{pmatrix} \cos\gamma\cos\theta & \sin\gamma & \cos\gamma\sin\theta \\ -\cos\theta\sin\gamma & \cos\theta & -\sin\theta\sin\gamma \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}
\end{aligned}$$

After the rotation, the local ellipsoidal distance metric points towards the observer throughout the whole galaxy sample. We use $\mathbf{A} = \mathbf{A}_\lambda(\gamma \;\; \theta)$ to define an anisotropic geometry estimator α_{A} for a given λ:

$$\alpha_\lambda(\mathbf{x}_i \;\; r) = \frac{2 \sum_{j=1}^{N} (\mathbf{d}_{ij}^{\mathrm{T}} \, \mathbf{A}_\lambda \, \mathbf{d}_{ij} \;\; r^2) \, \exp[-(\mathbf{d}_{ij}^{\mathrm{T}} \, \mathbf{A}_\lambda \, \mathbf{d}_{ij} \;\; r^2)]}{\sum_{j=1}^{N} \exp[-(\mathbf{d}_{ij}^{\mathrm{T}} \, \mathbf{A}_\lambda \, \mathbf{d}_{ij} \;\; r^2)]} \tag{3.28}$$

This equation is a more general definition of the local geometry than Equation (3.21) because the shape and orientation of the sampling volume are free parameters. The estimator can be adjusted and optimized to the underlying galaxy distribution. Yet, the physical interpretation of results may be ambiguous since the calculated geometries depend on the choice of those parameters included in the metric, which weights the distance to neighbor galaxies accordingly. Neighbor galaxies that are on the line-of-sight contribute more weights than the neighbor galaxies perpendicular to the line-of-sight. This non-Euclidean weighting of distances between galaxies might reduce the correlations between the geometry and galaxy properties, which are function of the Euclidean distance.

However, the observed galaxy distribution includes inherent redshift distortion that displace the location galaxies along the line-of-sight. This displacement can increase the distance between two neighbor galaxies with a similar evolution history, which leads to a reduced correlation of galaxy properties within the local environment. Because the anisotropic geometry estimator in Equation (3.28) adds more weight to the neighbor galaxies along the line-of-sight, the effect of the redshift distortions may be reduced.

In this work, we will use α_λ as an indicator for redshift space distortions. By adjusting the eigenvalues of the metric, those galaxy environment that are altered by the redshift space distortions can be classified. A motivation for the proper choice of the eigenvalues will given in Section 3.3. This classification can be helpful in the analysis of Fingers of God , clusters of galaxies that are elongated due to the redshift-space distortions.

3.2 Illustration with 3D Toy Models

After the definition of the geometry estimator α, we present an overview of possible results for α, which facilitates the interpretation for the measured geometries in the observed galaxy distribution of SDSS. We illustrate the values of α with the help of three dimensional toy models, artificially created point distributions with certain properties. Every toy model mimics a single isolated structure element of the observed large-scale galaxy distribution. The toy models allow to investigate the geometrical properties of the structure elements without the interference from neighboring structures. The volume surrounding the structure elements is filled with a random

point distribution with a mean density similar to the galaxy sample. This random distribution simulates the background signal from the surrounding galaxy distribution. The geometry α for the central point can be considered to be a typical value for each structure. By choosing a box size much larger than the sampling volumes, edge effects can be avoided.

The five toy models are generated by placing each of the five major structure elements, clusters, filaments, sheets, uniform regions and voids, in the centre of a box with the length $100\,h^{-1}$Mpc. The space around these structure elements is filled with a random point distribution with the mean density $0\ 0059\,h^3\mathrm{Mpc}^{-3}$, the mean density of the volume limited sample of the SDSS sample (Section 2.2).

The toy models were created according to the following specifications:

- The **cluster**-like structure was created by a uniform distribution of 50 points within a sphere radius equal $2.5\ h^{-1}$Mpc. The central point has the density contrast $\delta_8 = 3\ 9$ (top-hat filter, radius $8\,h^{-1}$Mpc).
- The **filament**-like structure was created by filling a cylinder (radius $1\,h^{-1}$Mpc, length $20\,h^{-1}$Mpc) with 50 randomly distributed points; its central point has a density of $\delta_8 = 2\ 5$.
- The **sheet**-like model was simulated with a flat cylinder (radius $15\,h^{-1}$Mpc, height $1\,h^{-1}$Mpc) filled by 100 points resulting in a density contrast of $\delta_8 = 1\ 9$.
- A **uniform** region was created by filling the whole box with a random distribution and adding a point to the centre ($\delta_8 = 0\ 088$).
- The **void**-like model for a void galaxy was created by a uniform distribution and removing all galaxies within a radius of $10\,h^{-1}$Mpc around the central galaxy ($\delta_8 = -0\ 92$).

The five toy models are shown in the five rows of Figure 3.2. Each panel in left and middle column shows a $10\,h^{-1}$Mpc slice in the x-y-plane and x-z-plane of the toy model box, respectively. The figures in right column show two geometries for various radii r estimated for the central point: the scaling index α and an approximation for α, a local slope $m(r)$ given by Equation (3.12):

$$m(r) = \frac{\mathrm{d}\ln N(r)}{\mathrm{d}\ln r} \approx \frac{\ln N(r + r_c) - \ln N(r - r_c)}{\ln(r + r_c) - \ln(r - r_c)} \tag{3.29}$$

where an arbitrary value of $r_c = 5\,h^{-1}$Mpc was chosen. In most models and for most radii, the geometry α and the local slope $m(r)$ have similar values. These examples indicate that the geometry α indeed estimates the slope of the function $\ln N(r)\ \ln r$, the change of number of points within a radius interval. The overall slight deviation and the large difference of the void toy model between $\alpha(r)$ and $m(r)$ can be both explained by two different weighting functions: a smooth Gaussian-like sampling sphere was used for the estimation of α, whereas the slope $m(r)$ was calculated with a top-hat filter.

In Figure 3.2, all toy models share a similar geometry behavior towards the lower and upper limit. For decreasing radii r 0, no neighbor galaxies are included in the shrinking sampling volume. The geometry of the center galaxy appears to be a point-like structure, $\alpha \sim 0$. For increasing radii $r > 10\,h^{-1}$Mpc , the large sampling volumes exceed the typical sizes of the structures. The features of the structures become marginal in the volume average and the geometry index approaches the value of uniformly distributed pattern, $\alpha \sim 3$, the dimension of the configuration space.

In addition to this observation, the extrema and slope changes of the function $\alpha(r)$ indicate rough approximations for the sizes of modeled structures in the centre of the box. The cluster model shows a clear geometry transition at the scale of the cluster radius of $\sim 2\ 5\ h^{-1}$Mpc. Below this scale, mostly points within the cluster are covered by the sampling volume, which were designed to be uniformly distributed. At the transition point of the cluster diameter, α reaches its lowest value due to the concentrated, point-like geometry. For increasing radii, the

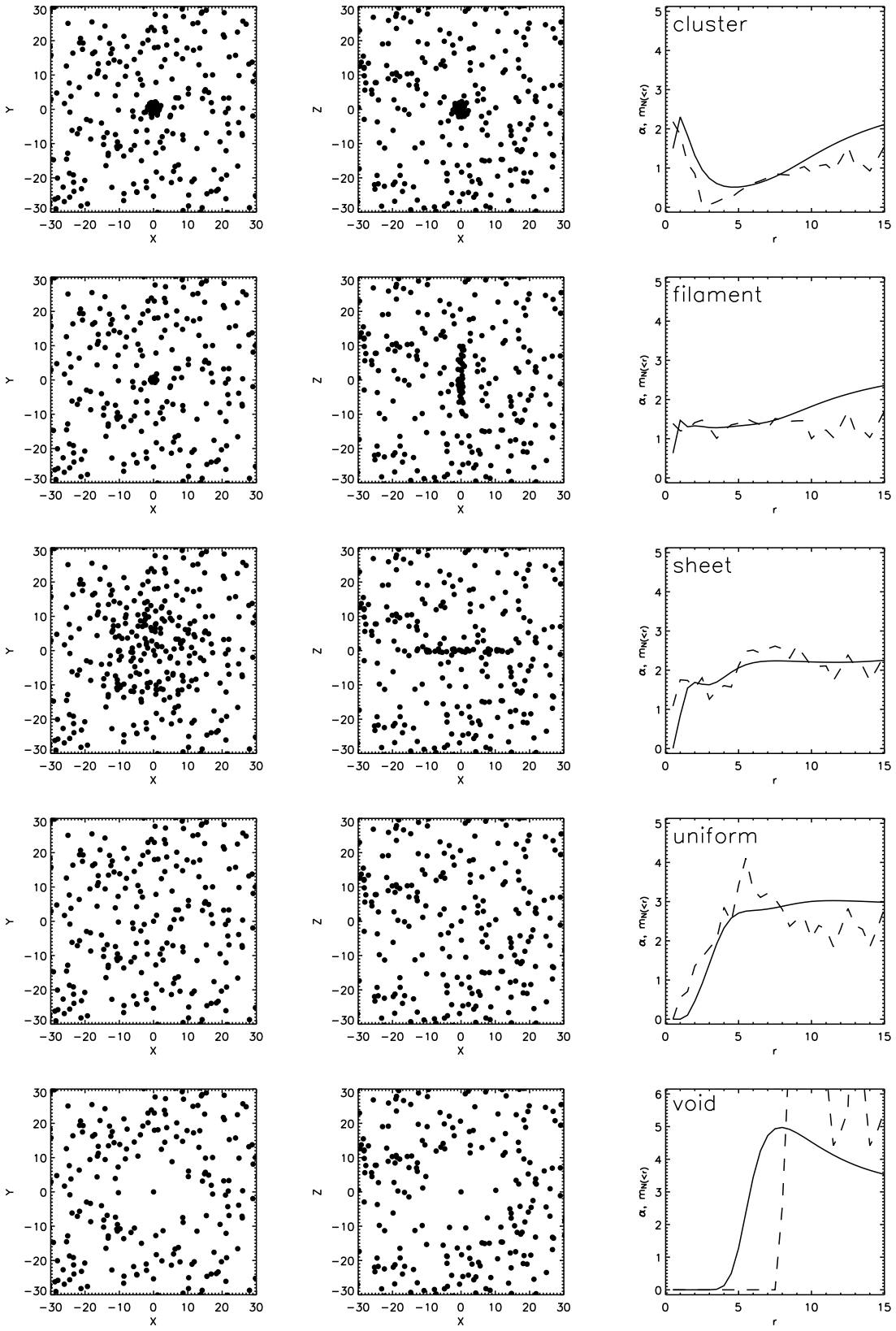

Figure 3.2: Five toy models of structure elements in the galaxy distribution and their corresponding α-values are shown (rows). Each panel in left and middle column shows a $10\,h^{-1}$Mpc slice in the x-y-plane and x-z-plane of the toy model box, respectively. The right column shows two geometry estimations for the central point as a function of the radius r: the scaling index α (solid line) and an approximation for α, a local slope $m(r)$ (dotted line, Equation 3.29).

cluster density decreases and the α-value increases. In the filament model, the central point of the filament has a constant, line-like geometry $\alpha \approx 1\,3$ for $r < 8\,h^{-1}$Mpc. Without the surrounding uniform distribution the geometry would be equal $\alpha = 1$; the points of the uniform distribution increase the denominator in Equation (3.12). For increasing radii, the geometry increases as well and the radius, for which the geometry increases, marks a rough estimate for the half length of the filament. The sheet model has a constant value of $\alpha \approx 2\,2$ for $r > 5\,h^{-1}$Mpc, which is similar to the dimension of this flat geometry. The uniform model show a transition at a similar radius, but has a geometry value of $\alpha \approx 3$, the dimensions of the configuration space. The central point of the uniform distribution has a point-like geometry for $r \lesssim 5\,h^{-1}$Mpc because the mean distance to the nearest neighbor is $\sim 3\,1\,h^{-1}$Mpc within this random distribution. The point in the center of the void in the last row of Figure 3.2 is by definition a point-like geometry, thus $\alpha = 0$ for $r < 5\,h^{-1}$Mpc . The geometry peaks for the radius range similar the size of the void and decreases for larger radii.

These qualitative results from the toy models indicate that the size of the sampling volume, here given by the choice the radius r, is crucial for the identification of an existing pattern in the local environment around a point distribution. The radius must be large enough to cover the typical size of the pattern in order to calculate a similar geometry classification as an observer would manually assign to. If the radius is smaller, the point distribution within the pattern dominates the classification. If the radius is larger, neighbor patterns that are part of the surrounding network of structures are included in the sampling volume. The resulting geometry is then an average over all the included patterns or corresponds to the global properties of the network.

The choice of the radius thus influences the discriminative power of the geometry estimator α. Two main ranges for the radius r emerged where the α distinguishes best between structures in the toy models: for $r < 3\,h^{-1}$Mpc and $r > 8\,h^{-1}$Mpc. Within each of those two intervals, the geometries for each structure prototype can be ordered in the following manner:

$$r < 3\,h^{-1}\text{Mpc}: \quad \alpha(\text{cluster}) > \alpha(\text{filament}) > \alpha(\text{sheet}) > \alpha(\text{uniform}) > \alpha(\text{void}) \tag{3.30}$$

$$r > 8\,h^{-1}\text{Mpc}: \quad \alpha(\text{cluster}) < \alpha(\text{filament}) < \alpha(\text{sheet}) < \alpha(\text{uniform}) < \alpha(\text{void}) \tag{3.31}$$

In the range $3\,h^{-1}\text{Mpc} < r < 8\,h^{-1}\text{Mpc}$, the α-r-relation shows different transitions for the different geometries, which depend on the local density of the considered point. In this range, two geometries can be distinguished, but no general interpretation for the α-values for all geometries can be given.

The presented toy models show that the different patterns of galaxy groups and alignments result in different local geometries that can be estimated and distinguished with the scaling index α. The geometries classify the local environments not only by measuring the abundance of neighbor galaxies, but also by including the information of their relative position. In addition to the local density, the local geometry seems to be a necessary description for a galaxy s environment.

3.3 Correction for the Redshift Space Distortions

3.3.1 General Effect of Redshift Distortions

The pattern of the large-scale structure of the Universe is assessed by measuring the three dimensional location of each galaxy in the sample: the two angle coordinates on the hemisphere and the radial distances to the observer. These distances are measured in an indirect way using the redshift z information of each galaxy, which is the result of the Hubble flow. For nearby galaxies ($z < 0\,1$), the radial, comoving distance r of a stellar system is given by

$$r = cz \;\; H_0 \tag{3.32}$$

Apart from the overall motion of the galaxy distribution due to the Hubble flow, stellar systems possess peculiar velocities $\mathbf{v}$ because they are attracted by gravitational potentials formed by

groups or clusters of galaxies. For example, the Local Group is attracted by the Virgo Cluster and both systems are moving towards the Great Attractor (Lynden-Bell et al. 1988; Bertschinger et al. 1990). These peculiar velocities overlap with the Hubble flow and alter the linear relation between the measured redshift and the proper distance of galaxies. The resulting distance s of a luminous object in redshift space can be written by

$$\mathbf{s} = \mathbf{r} + \hat{\mathbf{r}} \cdot \mathbf{v} \;\; H_0 \tag{3.33}$$

where r is the distance in real space. The term $v = \hat{\mathbf{r}} \cdot \mathbf{v}$ is the peculiar velocity, where the galaxy s velocity $\mathbf{v}$ is projected on the line-of-sight of an observer at $\hat{\mathbf{r}}$. Galaxies that move towards the observer are blue shifted and appear closer, whereas galaxies that move away appear redder and appear more distant. In general, these so-called redshift space distortions are systematic and unavoidable measurement errors, if the redshift of an astronomical object is used as a distance indicator. The most prominent consequence of redshift space distortions is the Fingers of God - effect where clusters appear elongated along the line of sight. In addition, filaments are stretched along the line of sight and loose their line-like shape.

For instance, a galaxy cluster at $z = 0\,1$ has a typical velocity dispersion of $\sigma_v = 1000$ km/s (e.g. Carlberg et al. 1996). Due to the redshift distortions, the shift in radial distance of galaxies within this cluster is $\mathbf{s} - \mathbf{r} \;\approx 10\,h^{-1}$Mpc . Assuming that this cluster has an isotropic, sphere-like volume in real space, the cluster appears as an ellipsoid with a fraction of the semi-axis of ~ 10 in redshift space.

In Figure 3.3, the effects of the redshift space distortions on the large-scale structure are illustrated. Both panels show a two-dimensional slice of a simulated galaxy distribution (Cole et al. 1998, Section 3.3.2) . The upper part of the figure shows the real space distribution, the lower part the corresponding redshift space distribution. The high-density regions in the real space slice, are clearly elongated towards the observer in the redshift space sample. In addition, thin filaments appear broader in redshift space due to the peculiar velocities. Since the location of galaxies can be changed by redshift space distortions, the classification of environments and local structures is influenced as well.

These qualitative effects of the redshift space distortions were illustrated by Zehavi et al. (2002), who estimated the clustering characteristics of different galaxy types. The redshift-space correlation function $\xi(s)$ in Figure 3.1 differs from the real-space correlation function $\xi(r)$ inferred from the projected correlation function $w(r_p)$ in Figure 3.4 in the expected sense: $\xi(s)$ is depressed on small scales by velocity dispersions and enhanced on larger scales by coherent flows. In addition, the slope of $\xi(s)$ is shallower and the correlation length is larger $s_0 > r_0$.

In principle, the anisotropy on larger scales has implications for the matter density Ω_m. With linear perturbation theory, a redshift-space distortion can be modeled with the real-space correlation function $\xi(r)$. This relation can be used to constrain the parameter β,

$$\beta \equiv \Omega_m^{0\,6} \;\; b \tag{3.34}$$

where the bias parameter b is the ratio of galaxy fluctuations to mass fluctuations (Kaiser 1987; Hamilton 1998). A comprehensive, quantitative analysis of redshift space distortions is limited because only linear approximation can be used and no analytical correction is available (Kaiser 1987; Hamilton 1998). An attempt to reconstruct the velocity field out of the observed galaxy distribution in order to correct the distortions was made by Bertschinger et al. (1990). In addition, simulations of the galaxy distribution can used to compare the effects of redshift space distortions on the statistical measures by creating real space and redshift space samples of the identical volume.

3.3.2 Estimating the Effect with Mock Catalogues

The effects of redshift distortions on statistical measures of the large-scale galaxy distribution can be obtained by numerical simulations, so called mock catalogs. These artificial created galaxy

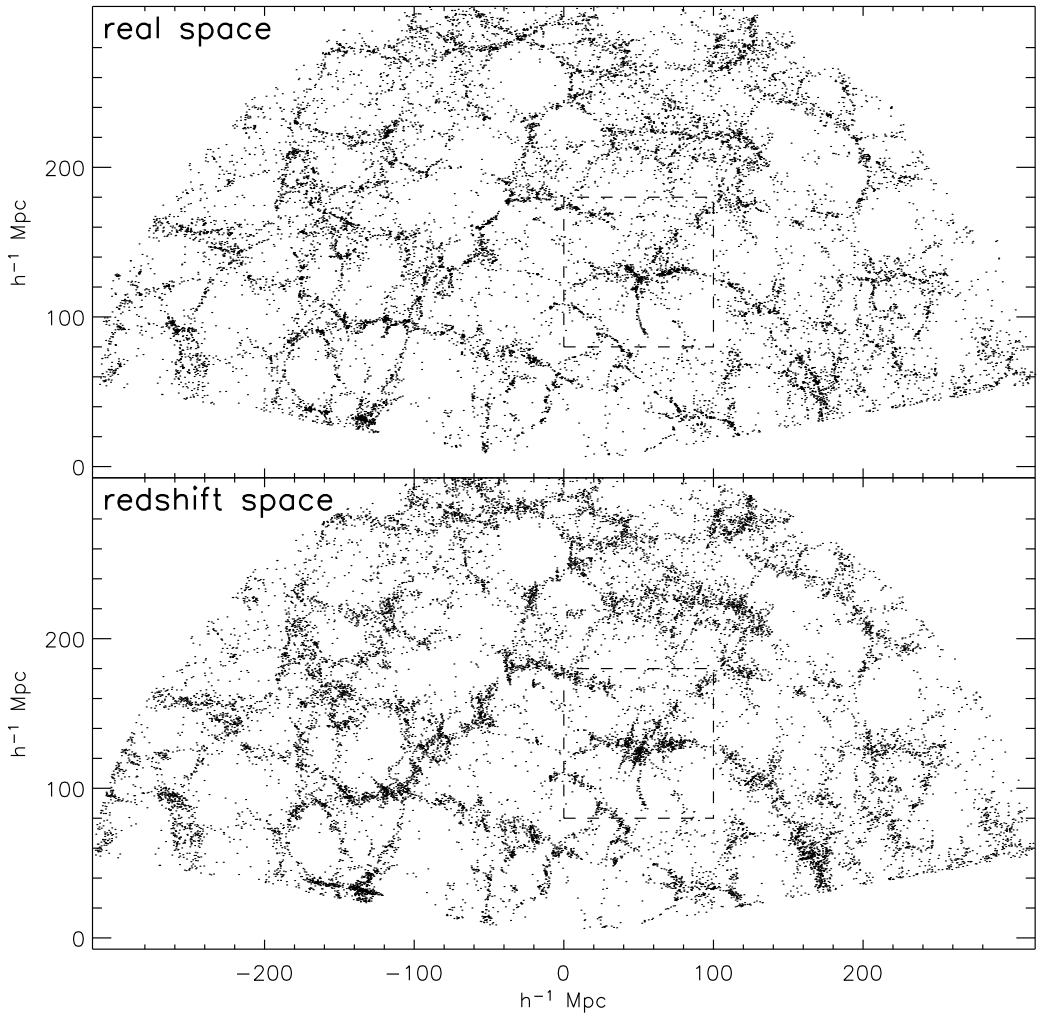

Figure 3.3: Comparison between real space (upper panel) and in redshift shift space (lower panel) clustering in with a mock catalogue. The slices are 4° thick in declination, flux limited and are limited to $z_{\max} = 0\ 11$. The region surrounded by the dashed line is magnified in Figure 3.5.

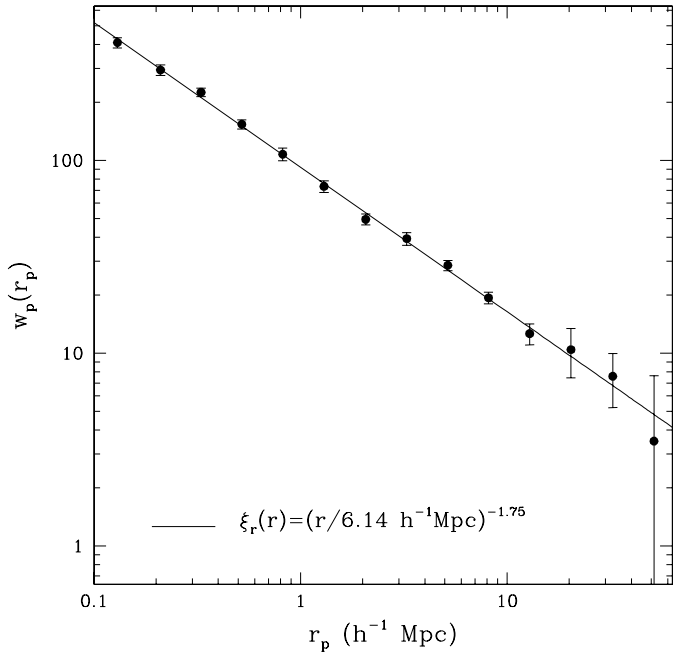

Figure 3.4: Projected correlation function $w_p(r_p)$ (filled circles) with a best fit for a power-law for w_p (solid line), which implies the denoted power-law for the real-space correlation function $\xi_r(r)$ (Zehavi et al. 2002).

distributions provide the clustering information about both the real space and the redshift space. By comparing the two distributions, the effect of the distortions on the geometry and density measures can be estimated in a statistical sense. The results of this investigation can indicate the typical scales and the magnitude of the effects, and can provide an approximation for the expected error for measures obtained from the observed galaxy distribution.

For this analysis, the mock catalogues described by Cole et al. (1998)[1] were used, which were designed to mimic the properties of the SDSS data, e.g the survey volume. Furthermore, the mock catalog includes a sophisticated bias scheme to identify galaxies in the simulated mass distribution. From the different data sets, we selected a galaxy distribution in a flat background cosmology with the matter density $\Omega = 0\ 3$ and a cosmological constant $\Lambda = 0\ 7$. More details about selected the mock catalog are given in the Appendix A.2. From the magnitude limited galaxy distribution, a volume limited sample was defined according to the only waveband that is provided by in the mock catalog, the b_J-band. Galaxies are included in this volume limited sample if they fall the absolute magnitude range $-22 < M_{b_J} < -18\ 9$ and have redshift limit of $z < 0\ 11$. The mean galaxy density of this sample results in $\bar{n} = 0\ 0057\, h^3\mathrm{Mpc}^{-3}$.

Since the peculiar velocities of the galaxies in the mock catalog are known, one can create two realizations of the identical galaxy distribution: a real space sample with the true positions of the galaxies, and a redshift space sample that mimics the observed galaxy distribution by including the redshift space distortions. For a selected position in the sample that corresponds to the observer, the radial velocities parallel to the line of sight are included in the determination of the distance to the observer by using Equation (3.33). The large-scale structure of the galaxy distribution in redshift space sample can then be compared with the real space galaxy sample without any redshift space distortions.

In Figure 3.3, the projection of two corresponding slices of the simulated galaxy distribution

[1] `http://star-www.dur.ac.uk/~cole/mocks/main.html`

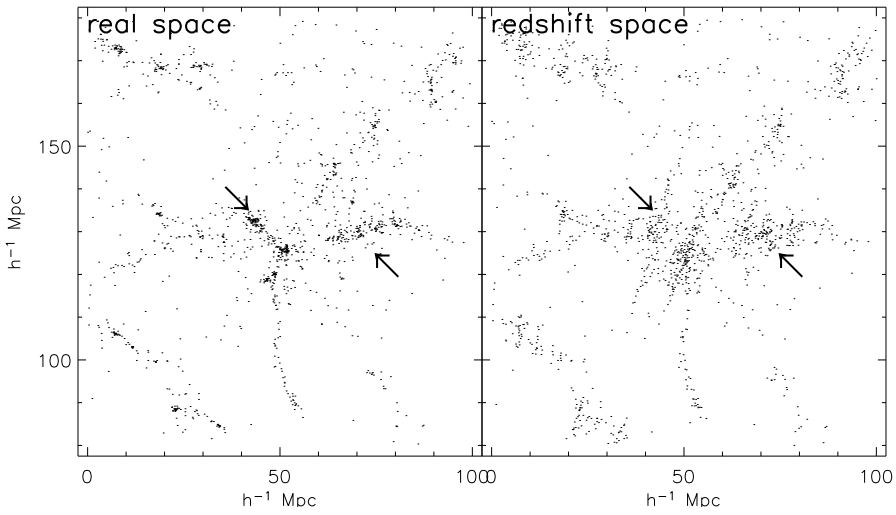

Figure 3.5: Comparison between real space (left panel) and redshift space clustering (right panel). The panel are two magnified areas of the corresponding, highlighted regions shown in Figure 3.3. The left panel shows a Finger-of-God prototype.

in the redshift space and the real space are shown. In the real space sample (upper panel), high-density regions are more point-like clustered, and filaments and void-like regions are more pronounced. In the redshift space (lower panel), the clusters are elongated along the line-of-sight to the observer, which is located in the origin of each panel. Galaxies are shifted away from the cluster s center and towards or away from the observer, depending on the galaxies peculiar velocity. The isotropic clustering in the real space appears more anisotropic in redshift space. Due to the same effect, the line-like structure of filaments becomes broader and appears fuzzier. However, on larger scales, the clustering of the galaxy distribution is very similar between in real space and redshift space samples. The overall structure with large void areas, separated with filaments and cluster-like areas is preserved.

The effects of the redshift space distortions on smaller scales are even more prominent in Figure 3.5, which consists of two magnified areas of the corresponding, highlighted regions shown in Figure 3.3 (dashed lines). The two arrows point in each panel mark two identical position in the two samples: a cluster region (left arrow) and a filament region (right arrow). At both locations, the galaxies in redshift space are shifted away from galaxy clusters or groups into the void-like regions. For the cluster region, the galaxies in the redshift space sample are considerably shifted within an elongated region, a prototype for a Finger-of-God , one of the most prominent redshift space distortion.

These illustrations indicate that there are differences in the galaxy clustering between the real space and the redshift space sample and that those differences are scale dependent. Any statistical measure that classifies the local environment of an observed galaxy dependent on the distance and location of surrounding galaxies will be affected by these distortions. This might lead to wrong classifications in observed data. However, not all galaxies might be affected by this error alike because the redshift space distortions depend on the scale of the volume of interest and on the environment in which a galaxy is located.

3.3.3 Scale Dependence

In order to investigate the influence of the redshift distortions on the geometry, the difference Δ_α between the geometries in the real space, α_r, and in the redshift sample, α_s, is calculated for each

radius	Δ_α		Δ_δ	
(h^{-1} Mpc)	μ	σ	μ	σ
2	0.172	0.441	-0.091	0.286
3	0.203	0.451	-0.048	0.269
5	0.124	0.408	0.002	0.202
8	0.040	0.295	0.020	0.131
10	0.009	0.235	0.025	0.103
15	-0.025	0.148	0.023	0.066

Table 3.1: The global redshift error distributions for the geometry α and the density δ.

galaxy and for a set of radii:

$$\Delta_\alpha = \alpha_\mathrm{s} - \alpha_\mathrm{r} \tag{3.35}$$

For the local density contrast, a corresponding error function Δ_δ can defined for the difference between δ_r and δ_s

$$\Delta_\delta = \delta_\mathrm{s} - \delta_\mathrm{r} \tag{3.36}$$

In order to avoid edge effects, galaxies are excluded if their distance to the simulation boundary is smaller than $2r$ (minus estimator, see Chapter 2).

In Figure 3.6, the effect of the redshift distortions on the geometry and density estimators is shown for different radii. The first row shows the distributions of the geometries α_r and α_s (pointed line) in real and redshift space, respectively. In the second row, the distribution of the geometry differences Δ_α is plotted and the dashed line corresponds to a fit to a Gaussian function $A\exp[-(\Delta_\alpha - \mu)^2\ (2\sigma^2)]$. The mean value μ and the standard deviation σ of Δ_α are given in Table 3.1. The rows c) and d) in Figure 3.6 are the corresponding results for the density δ and the density error Δ_δ.

The geometry differences are large at small scales $r < 5\,h^{-1}$Mpc, the typical cluster scale, where redshift space distortions are very prominent due to high velocity dispersions. On average, geometry values in redshift space are biased towards higher values ($\Delta_\alpha < 0$, Table 3.1) that indicates less pronounced structure elements. On larger scales ($r \geq 8\,h^{-1}$Mpc), the mean difference is close to zero and the standard deviation decreases. Thus, the effects on the geometry are less pronounced on the scales beyond the cluster scales and comparable with a randomly distributed noise. This scale dependence is consistent with the work of Fry & Gaztanaga (1994) who demonstrated with a correlation analysis that the flow of clustering power is shifted from small to the large scales due to redshift space distortions.

The corresponding results for local density differences are shown in Figure 3.6 c,d). On the smallest scales, the redshift space density is lower and the error function Δ_δ is biased towards negative values. In redshift space, the compact clusters appear elongated and the galaxy members are distributed over a larger volume. The lognormal fit does not model the error function. On large scales $r \geq 8\,h^{-1}$Mpc, the same effect results in higher estimated densities in redshift space, where galaxies clustering is reduced and less void galaxies are found. The overall error of the redshift effects is similar to a noise-like behavior and can be modeled by a lognormal fit.

Geometry and density are both affected by redshift space distortions on small scales and the absolute value of the error decreases with increasing radii. On cluster size scales, both properties are not measured accurately because the clusters are elongated to more filament-like structures. A correlation analysis of the both spaces confirmed that through the redshift space distortions correlation power is lost on scales smaller that cluster sizes; this power is transferred to larger scales (e.g. Kaiser 1987).

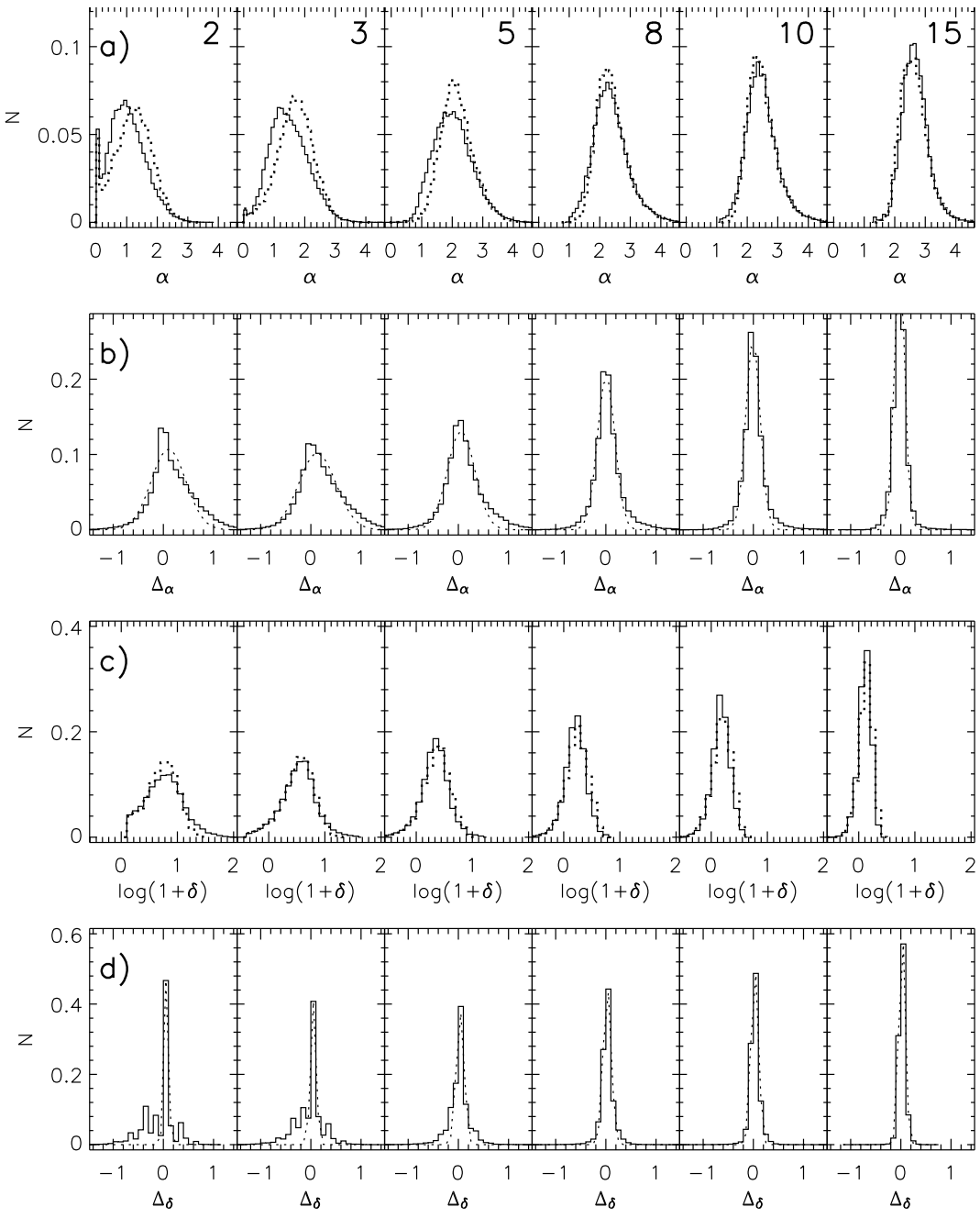

Figure 3.6: The effect of the redshift distortions on the geometry and density estimators is shown a set of radii (columns). Row a) shows the geometry distribution of $\alpha_{\rm r}$ and $\alpha_{\rm s}$ (pointed line) in real and redshift space, respectively. Row b) shows the distribution of the geometry difference Δ_α; the dashed line corresponds to a Gaussian fit. Row c) shows the density distribution of $\delta_{\rm r}$ and $\delta_{\rm s}$ (pointed line) in real and redshift space, respectively. Row d) shows the distribution of the density difference Δ_δ; the dashed line corresponds to a lognormal fit.

3.3.4 Dependence on the Environment

Apart from the scale dependence, the redshift space distortions can depend on the local environment, which can be classified with the geometry estimator α and the density contrast δ. It is interesting whether the distortions are a function of the surrounding large-scales structure, especially the observed redshift space structure. This would allow us to approximate the error with simulations, estimate it with observational data and apply error corrections. Both applications of this section will be incorporated in the observed relation in Chapter 4.

In Figure 3.7, the geometry error function Δ_α is plotted as a function of the geometry and density in real and redshift space. The upper row, Figure 3.7 a), visualizes the relation with the real space geometry α_r. On typical cluster scales $r \leq 5\,h^{-1}\mathrm{Mpc}$, the redshift error is anti-correlated. For $\alpha_\mathrm{r} \lesssim 1\,5$, the value of the real geometry is overestimated by the observed redshift geometry α_s. These galaxies in low density and void-like regions appear to be closer to structures like filaments and walls. Mock galaxies with $\alpha_\mathrm{r} \gtrsim 1\,5$ at this scales are found in clustered regions of space. The apparent diffusion of structures lead to a less homogenous environment and therefore to smaller values in geometry α_s. On larger scales $r \geq 8\,h^{-1}\mathrm{Mpc}$, a systematic relation between α_r and Δ_α is only seen for $\alpha \lesssim 2$. The geometry values for these scales correspond mainly to clusters and filaments in real space. They appear less structured in redshift space and are observed with a higher value $\alpha_\mathrm{s} > \alpha_\mathrm{r}$: clusters appear more filament-like, filaments appear wall-like. The correlation is not immanent for $\alpha_\mathrm{r} \gtrsim 2$. The values of Δ_α are randomly scattered around the zero point and show the stochastic nature in the redshift space distortions. In general, the redshift space distortions are most dramatic on the smaller scales, as already detected in the precedent section.

In Figure 3.7 b), the geometry distortions are plotted against the observable geometry α_s. Compared to the strong relation at the smallest scale $r = 2\,h^{-1}\mathrm{Mpc}$, the correlation decrease with increasing radius r. In addition, the error shows a more stochastic behavior with increasing radius and the magnitudes of the errors decrease. These results indicate that a linear approach to correct the redshift space distortions is not appropriate. However, for radii $r > 5\,h^{-1}\mathrm{Mpc}$, the distortions become small, $\Delta_\alpha < 0\,2$, and, on average, independent on the large-scale environment.

Further insight about the environment dependance is gained through the relations between the local density in real space δ_r and the geometry distortions Δ_α for all radii in Figure 3.7 c). On scales $r \leq 5\,h^{-1}\mathrm{Mpc}$, there is a dependence on the density of the environment, which can be separated into two regimes. For $\delta_\mathrm{r} \lesssim 3$, the mean value of the redshift distortions is $\Delta_\alpha \approx 0$ and only the variance increases with increasing density to maximum value. For the lowest densities $\delta_\mathrm{r} < 0$, the geometry distortions become negligible small, $\Delta_\alpha < 0\,1$. In the second regime $\delta_\mathrm{r} \gtrsim 3$, Δ_α correlates with the density and the variance in each density bin stays constant. The transition density between the two regimes and the typical scales correspond to the dense environments like galaxy clusters, where the Finger-of-God effect is a dominant geometry altering effect. The geometry distortions are biased towards positive values, which means that estimated redshift space geometry α_s is, on average, lower than the corresponding real space value. In addition, the magnitude of the error increases with the density on these scales. Similar results are found for dependance on the observable redshift space density δ_s in Figure 3.7 d) on the scales $r \leq 3\,h^{-1}\mathrm{Mpc}$, although the the geometry distortions depend less on the density δ_s. For the scales $r > 3\,h^{-1}\mathrm{Mpc}$, the geometry distortions are mostly scattered symmetrically around $\Delta_\alpha = 0$ similar to a random noise. Comparing the again the two densities on all scales, in redshift space less high density environments are estimated than in real space. Hence, with the observable redshift space density not all real space cluster regions with strong geometry distortions can be detected.

In summary, the redshift space distortions change the value of the geometry estimator on all scales and depends on the environment. The biggest error are most likely to occur on small scales ($r \leq 5\,h^{-1}\mathrm{Mpc}$) and in dense, cluster-like environments. However, there is a correlation between the observable redshift geometry and the real space geometry for those radii. For larger scales the errors decrease their magnitude but get stochastic with no significant dependence on the environment. We use these insights to model a real space geometry out of redshift space

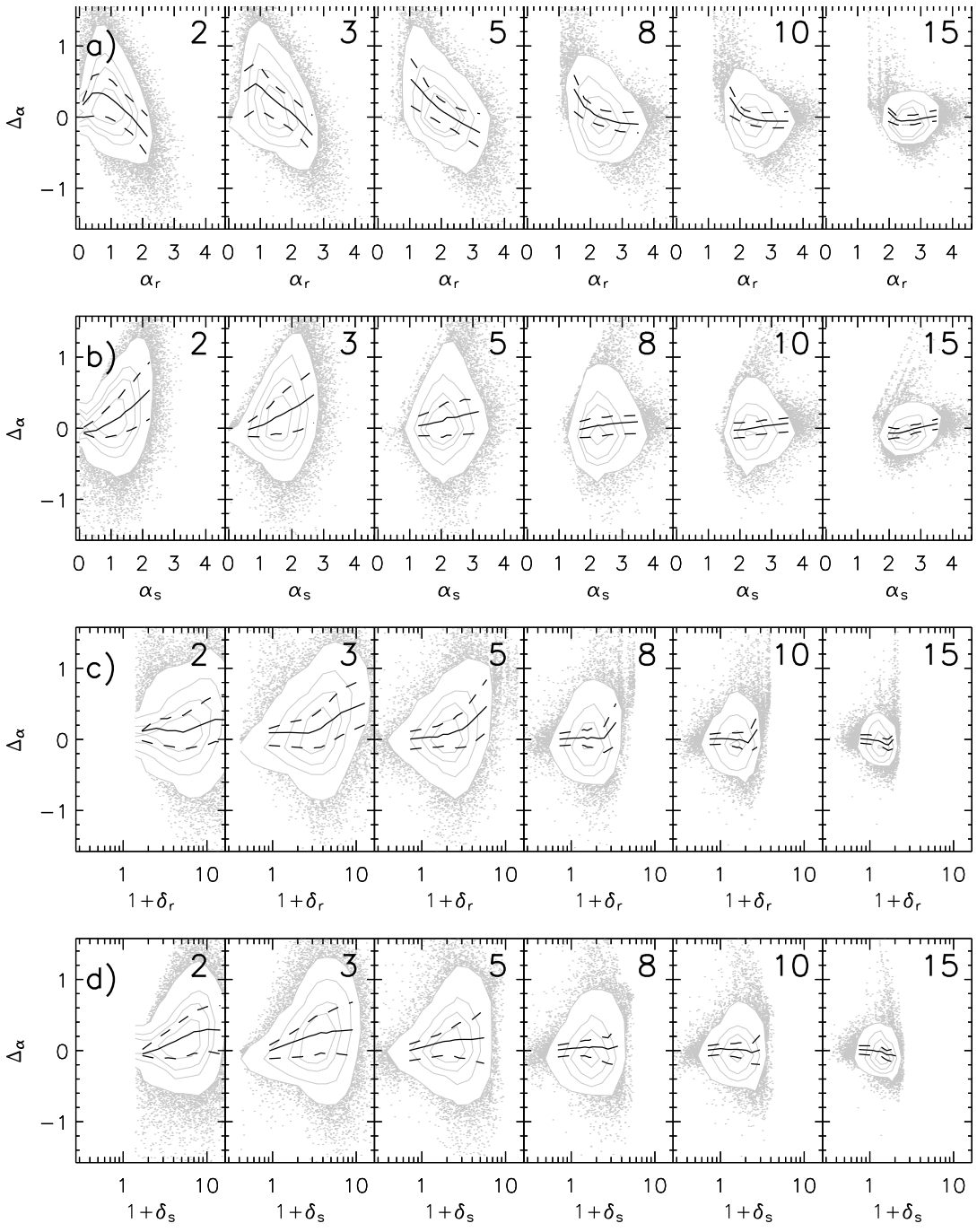

Figure 3.7: The difference between real space and in redshift shift space geometry Δ_α as a function of the real space geometry α_r (row a), redshift space geometry α_s (row b), real space density δ_r (row c), and the redshift space density δ_s (row d). The columns correspond to different radii (upper right corner). The point distributions in each sub-figure are represented by gray contour plots indicating levels of 10%, 30%, 50%, and 70% of the distribution; points outside these thresholds are plotted. In addition, the median (black solid line) and the 25th and 75th percentiles (dashed lines) are plotted, where all points were equally distributed to 10 bins.

radius (h^{-1} Mpc)	C_0	C_1	C_2	C_3
2	-0.118	0.439	-0.180	0.506
3	-0.364	0.362	-0.132	0.497
5	-0.477	0.231	-0.081	0.353
8	-0.407	0.141	-0.086	0.182
10	-0.355	0.111	-0.070	0.109
15	-0.258	0.080	-0.062	0.021

Table 3.2: The coefficients C_0 (corresponds to a constant), C_1 (geometry α_s), C_2 (density δ_s), and C_3 (anisotropic geometry $\alpha_s^{\rm A}$) for the linear regression fit.

properties. These corrections are created by different approaches.

3.3.5 Linear Regression Fit

The straightforward approach to predict the real space geometry α_r on the basis of redshift properties is to find correlations between the observable properties of the galaxies and the intrinsic error Δ_α due to the redshift space distortions which are gained through the simulations. Hence, we assume that Δ_α is a function of an arbitrary set of parameters that classify the environment in the redshift space. Here, we use the geometry $\alpha_{\rm s}$, the density $\delta_{\rm s}$, and the anisotropic geometry $\alpha_{\rm s}^{\rm A}$:

$$\Delta_\alpha = \Delta_\alpha(\alpha_{\rm s}\ \delta_{\rm s}\ \alpha_{\rm s}^{\rm A}) \tag{3.37}$$

This choice is motivated by the results from the previous section, where $\alpha_{\rm s}$ and $\delta_{\rm s}$ correlated with $\Delta_{\rm s}$. The anisotropic geometry $\alpha_{\rm s}^{\rm A}$ uses an ellipsoid sampling volume parallel to the line-of-sight and is sensitive to Finger-of-God -effect by construction. This approach aims to model mainly the distortion in high-density regions (clusters) and neglects the small effects for low-density regions (field galaxies). This is reasonable because cluster galaxies have significant higher peculiar velocities and, therefore, a higher radial displacement than field galaxies.

For a representative galaxy at the position $\mathbf{x}_i$, the error Δ_α at the radius r can be fitted by the following linear approach:

$$\Delta_\alpha(\mathbf{x}_i\ r) = \mathbf{s}_{ij}\mathbf{C}_j^{\rm T} \tag{3.38}$$

where

$$\mathbf{s}_{ij} = (1\ \alpha_{\rm s}(\mathbf{x}_i\ r)\ \delta_{\rm s}(\mathbf{x}_i\ r)\ \alpha_{\rm s}^{\rm A}(\mathbf{x}_i\ r)) \tag{3.39}$$

is a line vector including a constant, the redshift space variables geometry, density and anisotropic geometry, respectively. For scaling reasons, we use $\delta_s = \ln(1+\delta_s)$ to ensure that the values of variables are found in a similar range. The corresponding coefficients for a given radius are stored in the line vector

$$\mathbf{C}_j = (C_0\ C_1\ C_2\ C_3) \tag{3.40}$$

By applying a linear regression fit (Isobe et al. 1990; Akritas & Bershady 1996), these coefficients can be assessed for a given radii in Equation (3.38) and an estimated real space geometry $\alpha_{\rm r}^{\rm reg}$ can then be written as

$$\alpha_{\rm r}^{\rm reg} = \alpha_{\rm s} - \Delta_\alpha \tag{3.41}$$

where Δ_α includes the estimated coefficients $\mathbf{C}_j$. All galaxies in mock sample (Section 3.3.2) were used to assess the coefficients $\mathbf{C}_j$ and the results are presented in Table 3.2.

The dependence of the redshift error decrease with the sample radius as can be seen from the values of the coefficients. On small scales ($r \leq 5\,h^{-1}$Mpc), a strong linear dependence on $\alpha_{\rm s}$ (C_1) was found, as well a strong correlation with the anisotropic geometry estimator $\alpha_{\rm s}^{\rm A}$ (C_3). This

means that an increasing value of α corresponds to increased mean redshift error. On these scales, regions with galaxies cluster regions have the highest values of α_s and large redshift distortions. The same is valid for the anisotropic case, where a large value correspond to the cluster region where the redshift error is highest. Thus, C_4 has a the highest value on the smallest scale. The dependence on the density C_3 shows a negative correlation: the density in the s-space is underestimated with increasing density, because the volumes of clusters are increased through the elongation.

On larger scales ($r \geq 8\,h^{-1}$Mpc), the dependence on δ_s vanishs due to the decreasing influence of redshift space distortion in cluster regions. Not the peculiar velocities in high-density regions are dominating but the coherent large-scale galaxy flow into regions with lower density (e.g. voids). The anisotropic geometry $\alpha_\mathrm{s}^\mathrm{A}$, although still not negligible, is sensitive to cluster regions and looses importance, too. Whilst the isotropic counterpart α_s is still important up to $r = 10\,h^{-1}$Mpc , it achieves minor dependence at the largest scale like the density. Still present is the constant shift of the correction, the value of C_0, which is independent of the environment. This analysis again confirms the result, that the redshift space distortions have their greatest effect on the geometry on the smallest scales. Here, the effects are dependent on the environment, while on larger scales the effects become smaller and almost constant for all environments, i.e. that the redshift space distortions can be treated as a stochastic noise.

Before the estimated real-space geometry $\alpha_\mathrm{r}^\mathrm{reg}$ is compared with the true real-space geometry α_r in Section 3.3.8, more approaches to correct for the redshift space distortions are presented.

3.3.6 Principal Components Analysis (PCA)

After introducing the different parameters in order to characterize the environment of each galaxy, it is interesting to investigate the manifold of these variables in the parameter space. One can calculate the number of independent variables of the observed manifold to determine the intrinsic dimensionality of the problem. Furthermore, the relations between the observed quantities can be visualized. This can be accomplished by the principal component analysis (PCA), a well-known statistical tool for multivariate, astronomical data sets. For a set of correlated variables, the PCA technique identifies a substantially smaller set of uncorrelated variables. The analysis searches for the independent principal components, a few linear combinations of the original parameters that capture most of the information (variance) of the original parameters. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability. Further description and applications can be found in Deeming (1964), Brosche (1973), Murtagh & Heck (1987), Efstathiou & Fall (1984), and Folkes et al. (1999).

In general, a classification scheme for any set of objects can be established by assuming a set of distinct principal components $\mathbf{X}$. These components are defined in such a way that, for the observable quantities $\mathbf{s}$, there exist m relations $\mathbf{s} = \mathbf{s}(\mathbf{X})$. The problem of classification is then solved by inverting the transformation: $\mathbf{X} = \mathbf{X}(\mathbf{s})$. Since we possess no *a priori* knowledge about the functional form of $\mathbf{s}(\mathbf{X})$, we follow an empirical approach and derive a set of principal components $\mathbf{X}$ that are orthogonal linear combinations of the original variables, such that an inversion of $\mathbf{X}(\mathbf{s})$ is possible. The PCA is a systematic method which follows this approach.

Consider a sample of N galaxies ($i = 1 \qquad N$), each with M parameters ($j = 1 \qquad M$). If s_{ij} are the original measurements of the parameters, then construct the mean subtracted values:

$$s_{ij} = s_{ij} - \bar{s}_j \tag{3.42}$$

where $\bar{s}_j = \frac{1}{N}\sum_{i=1}^{N} s_{ij}$ is the sample mean. The covariance matix is then given by

$$C_{jk} = \frac{1}{N}\sum_{i=1}^{N} s_{ij}\, s_{ki}^\mathrm{T} \quad 1 \leq j \leq M \quad 1 \leq k \leq M \tag{3.43}$$

The matrix $\mathbf{C} = (C_{jk})$ is symmetric and can be reduced to a diagonal form:

$$\mathbf{U}^{\mathrm{T}}\mathbf{C}\mathbf{U} = \mathrm{diag}(\lambda_1 \quad \lambda_M) \tag{3.44}$$

where $\mathbf{U} = \begin{matrix}\mathbf{u}_1^{\mathrm{T}} & \mathbf{u}_M^{\mathrm{T}}\end{matrix}$ is the matrix in which the jth column is the eigenvector $\mathbf{u}_j$ of $\mathbf{C}$ corresponding to the eigenvalue λ_j. The axis along which the variance is maximal has the largest eigenvalue λ_1. It is convenient to sort the eigenvectors by their eigenvalues in decreasing order. With matrix $\mathbf{U}$ of all eigenvectors a new set of varibles $\mathbf{X}$ can be constructed that are orthogonal to each other:

$$\mathbf{X} = \mathbf{s}\mathbf{U} \tag{3.45}$$

With this linear transformation, we can construct a set of variables X_j that are orthogonal linear combinations of the observed set of parameters s_j. Those variables $\mathbf{X}$ are the principal components of the observation. The eigenvalue λ_j is then the independent variance of the data in the directions of the principal components X_j.

The inversion the transformation is given by

$$\mathbf{s} = \mathbf{X}\mathbf{U}^{\mathrm{T}} \tag{3.46}$$

since $\mathbf{U}$ is orthogonal by definition ($\mathbf{U}\mathbf{U}^{\mathrm{T}} = \mathbf{1}$). Because the principal components with the small eigenvalues are poorly constrained, it is reasonalbe to reduce the set variables the axis with the highest variances. Using only P principal components ($P < M$), the approximated variables $\tilde{\mathbf{s}}$ are given by

$$\tilde{\mathbf{s}} = \mathbf{X}\mathbf{U}_P^{\mathrm{T}} \tag{3.47}$$

where the columns of the matrix $\mathbf{U}_P = \begin{matrix}\mathbf{u}_1^{\mathrm{T}} & \mathbf{u}_P^{\mathrm{T}}\end{matrix}$ consists of the reduced set of principal components $\mathbf{u}_j$.

Note, that the results of the PCA method depend on the scaling of the parameters. The principal components are not unique and depend on the relative scaling of the variables s_i, as well of the their functional form, e.g. whether we use $\ln(1+\delta)$ or δ for the calculation. The physical interpretation of the principal components must be done carefully. Here, we use the PCA as tool in order to investigate whether the observational points $\mathbf{s}$ lie within a subspace of the M-dimensional hyperspace. Thus, we look for degeneracies among the physical parameters which may be unresolved by the observation.

We computed the eigenvalues and the principal components for the parameters listed in Table 3.3. For each radii, the four principal components are sorted by the their eigenvalues and are given in the four columns on the right side, and the corresponding eigenvalue λ, the fractional variance (FV), and the cumulative fractional variance (CFV) in the three columns on the left. The four elements of a principal components assign weights, so-called loading factors , to the four variables geometry α_{s}, local density $\delta_{\mathrm{s}} = \ln(1+\delta_{\mathrm{s}})$, anisotropic geometry $\alpha_{\mathrm{s}}^{\mathrm{A}}$, and geometry distortions Δ_α. These weights reflect the correlation and linkage of a variable with a given principal component.

For the smallest scale $r = 2\,h^{-1}$Mpc, the first principal component PC1 explains 67.8% of the variance in the data set, while first two PCs together explain 85.8% and first three PCs jointly explain 96.9% of the variance. Because PC1 depends mostly on the geometry α_{s} and the local density δ_{s}, it can be referred to as the component of the cluster environment. The geometry distortions Δ_α have a large weight on PC2 with FV of 18%, which means that this variable contributes to the considerable amount of information that is not explained by the other variables. The anisotropic geometry $\alpha_{\mathrm{s}}^{\mathrm{A}}$ has a strong linkage to PC4 with the low fractional variance FV=3.1%. Similar results were found for the scale $r = 3\,h^{-1}$Mpc. Both, PC1 and PC2, are linked with the geometry and the local density as well, but together only explain 77.8% of the variance.

On the scales $r = 5\,h^{-1}$Mpc and $r = 8\,h^{-1}$Mpc, the results are less conclusive. For both radii, PC1 and PC2 explain less of the variance contained in the data set, and both PCs are linked with geometry, density, and anisotropic geometry. The characteristics of the cluster environment

	α_s	δ_s	α_s^A	Δ_α	λ	FV [%]	CFV [%]
$r = 2\,h^{-1}$Mpc							
PC1	0.679	0.687	0.059	0.253	0.662	67.8	67.8
PC2	0.094	-0.422	-0.030	0.901	0.176	18.0	85.8
PC3	0.663	-0.499	-0.458	-0.318	0.109	11.2	96.9
PC4	0.301	-0.318	0.886	-0.151	0.030	3.1	100.0
$r = 3\,h^{-1}$Mpc							
PC1	0.612	0.750	0.101	0.229	0.608	57.0	57.0
PC2	-0.628	0.564	0.408	-0.347	0.222	20.8	77.8
PC3	-0.299	-0.073	0.381	0.872	0.186	17.4	95.2
PC4	0.375	-0.338	0.823	-0.259	0.051	4.8	100.0
$r = 5\,h^{-1}$Mpc							
PC1	0.127	-0.788	-0.573	-0.184	0.433	42.2	42.2
PC2	0.895	0.317	-0.282	0.136	0.330	32.1	74.3
PC3	0.128	-0.419	0.336	0.834	0.180	17.5	91.8
PC4	0.408	-0.321	0.692	-0.503	0.084	8.2	100.0
$r = 8\,h^{-1}$Mpc							
PC1	0.493	-0.441	-0.749	-0.050	0.496	54.5	54.5
PC2	-0.820	-0.478	-0.245	-0.196	0.207	22.7	77.2
PC3	0.183	-0.738	0.529	0.375	0.136	14.9	92.1
PC4	0.226	-0.178	0.314	-0.905	0.072	7.9	100.0
$r = 10\,h^{-1}$Mpc							
PC1	0.488	-0.373	-0.789	-0.014	0.479	59.2	59.2
PC2	-0.858	-0.156	-0.453	-0.185	0.167	20.7	79.9
PC3	0.069	0.909	-0.384	-0.149	0.113	13.9	93.8
PC4	0.145	-0.104	0.157	-0.971	0.050	6.2	100.0
$r = 15\,h^{-1}$Mpc							
PC1	-0.404	0.253	0.878	-0.029	0.381	59.9	59.9
PC2	0.848	-0.235	0.462	0.109	0.175	27.6	87.5
PC3	0.330	0.936	-0.119	-0.040	0.059	9.3	96.8
PC4	-0.091	0.071	-0.030	0.993	0.020	3.2	100.0

Table 3.3: The four principal components (PC) for each reach radius are shown. The four columns on the left of each line consists of the principal components components (loading factors) for the variables geometry α_s, local density $\delta_s = \ln(1 + \delta_s)$, anisotropic geometry α_s^A, and geometry distortions Δ_α. The three columns on the right show the eigenvalue λ, fractional variance (FV) and cumulative FV (CFV) for the corresponding principal components.

radius (h^{-1} Mpc)	C_0	C_1	C_2	C_3
2	0.633	2.072	-2.193	6.183
3	-1.471	1.497	-1.343	3.267
5	-1.736	0.857	-0.836	1.465
8	-0.693	0.291	-0.605	0.479
10	-0.478	0.166	-0.291	0.204
15	-0.277	0.089	-0.108	0.029

Table 3.4: The coefficients for the transformation from the coordinate system defined by the principal components to the original coordinate systems for a given radius (rows).

become less dominant and the variety of all environment reduces the correlations within the variables. Thus, more variables (information) are needed to explain the full variance and no variable is obsolete. This leads to the high observed value of FV (up to 8.2%) for PC4.

On the scales $r = 10\,h^{-1}$Mpc and $r = 15\,h^{-1}$Mpc, the anisotropic geometry has the highest absolute weight for the PC1. Together with high linkage of the geometry α_s, PC1 and PC2 refer to the detection of the elongated cluster regions, the Finger-of-Gods , which are the main and dominant effect of the redshift distortions and have a distinct geometry. On these scales, the density provides only few information and has a strong linkage with PC3. The large sampling volume underestimated small high-density regions, where the redshift distortion are large; the sensitivity to clusters is reduced and the changes in the galaxy environment cannot be detected. The variable Δ_α is linked with PC4 and contributes the least information.

In summary, the results in Table 3.3 indicate that the PCA analysis works best on the small and the large scales of the used set of radii where the data set could be reduced to three independent principal components which together still explain more than 95% of the variance. On small scales, PC1 refers to cluster regions, the origin of geometry distortions, that still have a high density in redshift space and can be characterized by geometry and density estimators. On the large scales, PC2 refers to the main redshift space distortions, the Finger-of-God, which have a high linkage with the anisotropic geometry. For completeness, the coefficients for the transformation from the PC coordinate system into the original coordinate systems are given in Table 3.4.

3.3.7 Environment Fit

A third approach to model the redshift space distortions can be created by using the mock catalog as a reference galaxy distribution that includes all possible galaxy environments. One can assume that such a library includes all possible galaxy environments which have the full information of the effect of the redshift distortions because the velocities of the mock galaxies are known. By finding the closest match between an observed galaxy and a simulated library galaxy in redshift space, one can use the real space properties of the simulated galaxy as an estimate for the real space property of the observed galaxy.

Here, we use observed environment properties of the SDSS galaxies which define their surrounding in the redshift space (s-space) and find the best fitting s-space environment in the simulated library. Then, we assume that the real space (r-space) properties between observation and simulation are *a priori* similar which is justified by the construction of the mock catalog (see Chapter 3.3.2). The r-space geometry of the best fit mock galaxy in the library can then be used for an estimation of the r-space geometry of the observed galaxy.

In addition of the construction of the mock catalog, the validity of this approach depends on the best possible classification of the s-space environment. In order to compare the results with the other two approaches, we use again use the arbitrary set of three properties: the isotropic

geometry $\alpha_{\rm s}$, the local density with $\delta_{\rm s}$ and the anisotropic geometry $\alpha_{\rm s}$. All three together classify the geometry of the galaxy distribution and are sensitive to the redshift space distortions. We define the fitting function $\xi_{\,m}$ of a mock galaxy m and an representative observed galaxy as the sum over squared difference of the used environment properties at the radii r_i:

$$\xi_{\,m} = \sum_{i=1}^{6} \Delta_{\alpha}^{2}(r_i) + \Delta_{\delta}^{2}(r_i) + \Delta_{\alpha^{\rm A}}^{2}(r_i) \tag{3.48}$$

where

$$\Delta_{\alpha}(r_i) = \alpha_{m\,{\rm s}}(r_i) - \alpha_{\rm s}(r_i) \tag{3.49}$$

is the geometry difference between the value of the mock galaxy $\alpha_{m\,{\rm s}}$ and the observed value $\alpha_{\rm s}$ at the radius r_i. The used radii r_i are identical with the set of radii of the previous analysis: r_i $(h^{-1}{\rm Mpc}) =$ 2 3 5 8 10 15 . The density difference $\Delta_{\delta} = \ln(1+\delta_{m\,{\rm s}}) - \ln(1+\delta_{\rm s})$ and the discrepancy in the anisotropic geometry $\Delta_{\alpha^{\rm A}} = \alpha_{m\,{\rm s}}^{\rm A} - \alpha_{\rm s}^{\rm A}$ are defined in the corresponding manner. The function $\xi_{\,m}$ provides a quantitative decision to find the mock galaxy m that is most similar to the observed galaxy in s-space, $\xi_{\,m} = \min(\xi_{\,m})$, in order to use the corresponding r-space geometry:

$$\alpha_{\rm r}^{\rm lib} \sim \alpha_{m'\,{\rm r}\;\gamma_{m'}=\min(\gamma_m)} \tag{3.50}$$

In order to reduce the sensitivity to outliers and to increase the statistical stability, not only one mock galaxy with the minimal value of $\xi_{\,m'}$ is taken, but the average of a set of 10 mock galaxies with the 10 smallest values of $\xi_{\,m}$ is used for each observed galaxy. This mean real space geometry of these mock galaxies is chosen to be the estimated r-space geometry for an observed galaxy $\alpha_{\rm r}^{\rm lib}$:

$$\alpha_{\rm r}^{\rm lib} = \sum_{i=1}^{10} \alpha_{m_i\,{\rm r}}\;\;10 \tag{3.51}$$

To evaluate the performance of this approach, the mock catalog is divided in a test sample with 2000 randomly selected galaxies and a reference sample including the remaining galaxies. For the galaxies in the test sample, the environment fit that is described in Equation (3.48) was performed. For the test sample, the mean and standard deviation of the minimum value is $\xi_{\,m'} = 0\;19 \pm 0\;12$, of the 10th-smallest value $\xi_{\,m\;10{\rm th}} = 0\;46 \pm 0\;20$, and the mean of all the 10 smallest values $\xi_{\,m\;10} = 0\;36 \pm 0\;16$. In comparison, the overall mean value was, on average, significantly larger, $\xi_{\,m} = 9\;5 \pm 4\;6$. This confirms the trivial fact that galaxies reside in different environments. By applying density and geometry measures, those environments can be classified into structure groups, e.g. the group of cluster members. For each galaxy, there exists another galaxy in the sample, which was assigned to the same structure group and, thus, has similar structure and presumably similar physical properties. This will be discussed further in Chapter 4 and 5.

After the calculation of the mapping, the true r-space geometry α_r can then be compared with the estimated r-space geometry $\alpha_{\rm r}^{\rm lib}$ by calculating a error function similar to Equation (3.35):

$$\Delta_{\alpha}^{\rm lib} = \alpha_{\rm r}^{\rm lib} - \alpha_r \tag{3.52}$$

In Table 3.5, the test sample s mean and standard deviation of $\Delta_{\alpha}^{\rm lib}$ are shown. In addition, the results are visualized in Figure 3.8. The following section will discuss the performance and compare it with the results of the other two approaches which were presented in the previous sections.

3.3.8 Evaluation of the Corrections

After presenting the different approaches to estimate the value of the r-space geometry from observable s-space properties, the performance of the different approaches are evaluated using the data from the mock catalog. The test and reference sample described in the previous section were used for this analysis. For each galaxy in the test sample, the true geometry difference Δ_{α}

radius	Δ_α		$\Delta_\alpha^{\mathrm{reg}}$		$\Delta_\alpha^{\mathrm{lib}}$		$\Delta_\alpha^{\mathrm{pca}}$	
$(h^{-1}\,\mathrm{Mpc})$	μ	σ	μ	σ	μ	σ	μ	σ
2	0.182	0.445	0.020	0.395	0.026	0.371	0.071	1.335
3	0.211	0.453	0.013	0.408	0.016	0.384	0.076	1.068
5	0.122	0.412	-0.000	0.388	0.008	0.376	-0.040	0.906
8	0.035	0.308	-0.001	0.295	0.010	0.278	-0.107	0.752
10	0.003	0.245	-0.001	0.238	0.005	0.227	-0.121	0.670
15	-0.033	0.150	0.003	0.146	0.002	0.159	-0.143	0.549

Table 3.5: The mean μ and standard deviation σ of the true and the estimated geometry differences are given for the different radii (rows) and different correction approaches.

was compared with the estimated geometry differences $\Delta_\alpha^{\mathrm{reg}}$, $\Delta_\alpha^{\mathrm{lib}}$, and $\Delta_\alpha^{\mathrm{pca}}$, which were assessed by the regression fit, the library fit, and the PCA analysis, respectively. All the estimated error function are defined similar to Equation (3.52).

In Table 3.5, the mean μ and standard deviation σ of the true and the estimated geometry differences are given for the different radii. On small scales, the true geometry difference is biased towards positive values and has a large scatter. This indicate that the redshift shift space geometry is overestimated on average. For scales larger than $r = 3\,h^{-1}$Mpc, the bias and the scatter decreases for increasing radii because the effect of redshift distortions becomes stochastic noise with $\mu \approx 0$ on those scales (see Section 3.3.3).

For the error functions $\Delta_\alpha^{\mathrm{reg}}$ and $\Delta_\alpha^{\mathrm{lib}}$, similar results were found. On all scales, the mean value is close to zero, $\mu \approx 0$, and the scatter σ decreases with increasing radius. Compared to Δ_α, μ is only marginally biased on the small scales $r \leq 5\,h^{-1}$Mpc, whereas the scatter σ is slightly reduced or similar between the true and the two estimated error functions on all scales. The standard deviation of $\Delta_\alpha^{\mathrm{reg}}$ is 12% smaller at the smallest scale and 5% at the largest. The library fit $\Delta_\alpha^{\mathrm{lib}}$ performed better on almost all scales, e.g. 17% less scattering at the radius $r = 2\,h^{-1}$Mpc; however, not for the largest radii, where a 8% higher standard deviation was observed. Thus, the two corrections are able to correct for the generall trend of the redshift space distortion to overestimate the geometry at small scales but have no significant effect on the larger scales ($r \geq 8\,h^{-1}$Mpc). Especially the library fit indicates that there is still information contained in the redshift environment from which an acceptable real space geometry can be estimated.

The PCA approach does not estimate acceptable geometry corrections, although the bias on small scales is slightly reduced. On all scales, the error function $\Delta_\alpha^{\mathrm{pca}}$ has a large scatter (up to $\sigma = 1\,335$ for $r = 2\,h^{-1}$Mpc) which is considerable higher than for the true geometry difference. In addition, the PCA approach induces a negative bias on the large scales which is not present in the true geometry difference nor the other two estimated geometry differences. This bad performance of the PCA approach may be explained by the non-Gaussian distribution of the used data. The PCA method assumes that each variable is distributed close to a Gaussian function. If this assumptions is not fulfilled, the calculation of the principal components is sensitive to outliers and non-Gaussian tails of the data variables. In Figure 3.7, those non-Gaussian properties are prominent on all scales for relations between geometry, density and geometry difference. Although the PCA corrections cannot be used, the method is still useful for a qualitative understanding of the results (see Section 3.3.6).

In addition to performance of the corrections on each scale, it is interesting to investigate how the estimated geometry difference depend on the galaxy s environment. It was shown in Section 3.3.4 that the redshift space distortion depend on the structure element a galaxy is embedded in and one can presume that the corrections show a similar behavior. In Figure 3.8, the four error distributions are plotted as a function of the true r-space geometry α_r.

Again, the functions $\Delta_\alpha^{\mathrm{reg}}$ and $\Delta_\alpha^{\mathrm{lib}}$ (Figure 3.8 b, c, respectively) show a similar behavior. Both

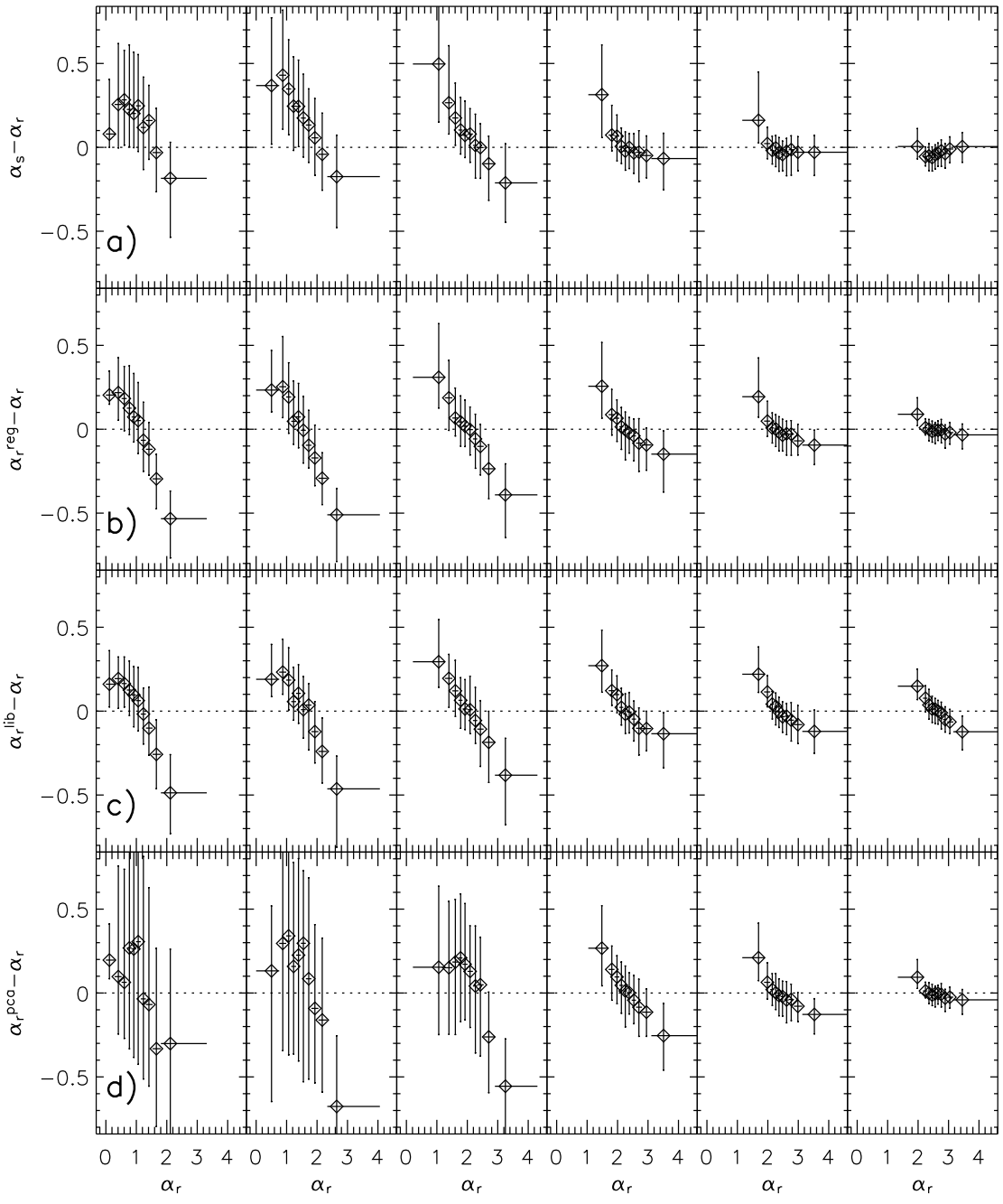

Figure 3.8: The true geometry difference $\Delta_\alpha = \alpha_s - \alpha_r$ as a function of the real space geometry α_r (panel a)) for the six different radii (columns). The estimated geometry differences $\Delta_\alpha^{\rm reg} = \alpha_r^{\rm reg} - \alpha_r$, $\Delta_\alpha^{\rm lib} = \alpha_r^{\rm lib} - \alpha_r$, and $\Delta_\alpha^{\rm pca} = \alpha_r^{\rm pca} - \alpha_r$ are presented in the panels b), c), and d), respectively. The data is divided into 10 bins which includes the same number of data points. The abscissa of each data bin is marked with its mean value; the horizontal error bar indicates the bin s half distance to the neighbor bin on each side. The ordinate of each bin is marked with its median value; the lower and upper value of the vertical error bars correspond to the 25% and 75% percentile of the data, respectively.

preserve the overall shape of the of the true geometry difference Δ_α and only apply some minor corrections. On the small scales, both corrections have a smaller bias and a smaller scatter for small α_r-values, $\alpha_r < 1\,5$ ($r = 2\,h^{-1}$Mpc) and $\alpha_r < 2$ ($r = 3\,5\ h^{-1}$Mpc), but assign a larger negative error for the larger α_r-values (e.g. cluster regions). On the larger scales, both corrections are slightly biased towards negative values for $\alpha_r > 2$. In addition, on the largest scale $r = 15\,h^{-1}$Mpc, both corrections have a positive bias for $\alpha_r < 2$, which is not prominent in the true geometry difference, and $\Delta_\alpha^{\rm lib}$ is slightly more biased for larger bins $\alpha_r > 2$. In contrast, the function $\Delta_\alpha^{\rm pca}$ (Figure 3.8 d) has a huge scatter on the small scales $r \leq 5\,h^{-1}$Mpc, but, compared to the other two corrections, a similar behavior for the larger scales, where the assumptions of the PCA method seem to hold for the used data set. In general, those plots suggest that the best corrections of the redshift space distortions are limited due to the fact that the distortions are nonlinear and not reversible.

Because the r-space geometry is not available in the later analysis of the observed data, the dependence of the observable s-space geometry is investigate in Figure 3.9 as well. On the small scales $r \leq 5\,h^{-1}$Mpc, Δ_α (panel a) correlates with the geometry $\alpha_{\rm s}$, but for all corrections (panels b, c, d) the mean values in each bin are close to independent from $\alpha_{\rm s}$. This indicates that the corrections work by averaging out the mean distortions based on redshift space properties. For $\Delta_\alpha^{\rm reg}$ and $\Delta_\alpha^{\rm lib}$, only the standard deviation correlates mildly with the small scales geometry of the environment. At these scales, larger geometry values correspond to the cluster environments where the high peculiar velocity dispersion increases the variation of redshift space distortions. Again, $\Delta_\alpha^{\rm pca}$ has a high scatter on the small scale. For $r \geq 8\,h^{-1}$Mpc, all panels in Figure 3.9 show similar behaviour in terms of bias and scatter.

In conclusion, the redshift space distortion are present on all scales with different effects. The most prominent effect was found in the cluster environments and on scales similar to the cluster dimensions. However, the redshift space geometry is still a good approximation for the true real space geometry. In estimating the real space geometry with mock catalogs, the environment dependent bias and mean deviation from the true r-space value can be reduced by at most 17%.

3.4 First Results for the SDSS Galaxies

After the definition of the geometry estimator α and the systematical redshift space distortions, first results for the observed SDSS galaxy sample are presented in order to evaluate the significance of the results. In Chapter 4, these results are used to investigate the relations between the geometry and others physical properties galaxies.

3.4.1 3D Results

The distribution function of the s-space geometry α for the SDSS galaxy sample described in Chapter 2 are given in Figure 3.10 for different radii (in the following discussion, α without the any index denotes the observed, redshift geometry). The histograms include a broad spectrum of geometries which correspond to different environments that are present in the galaxy sample. Not only high density galaxies (cluster regions) and low density galaxies (void-like regions) are found, but in addition, the α spectrum indicates that filament-like regions ($\alpha \approx 1$) and wall-like regions ($\alpha \approx 2$) are present in the galaxy distribution. While the mean value of the geometry distributions increase with the radius, the variation is almost constant. Note, that equal values of α have a different meaning for different scales and can correspond to different environments (see Section 3.2).

For the radii $r \leq 3\,h^{-1}$Mpc , the smallest values ($\alpha < 1$) were calculated for galaxies in underdense, void-like regions. Here, the distance between neighbor galaxies is large and only a few galaxies are located within the sampling volume. Those galaxies are not connected to any small-scale structure and have a point-like value of $\alpha \approx 0$. The higher values of the distribution correspond to higher density regions where the clustering of the neighbor galaxies appears

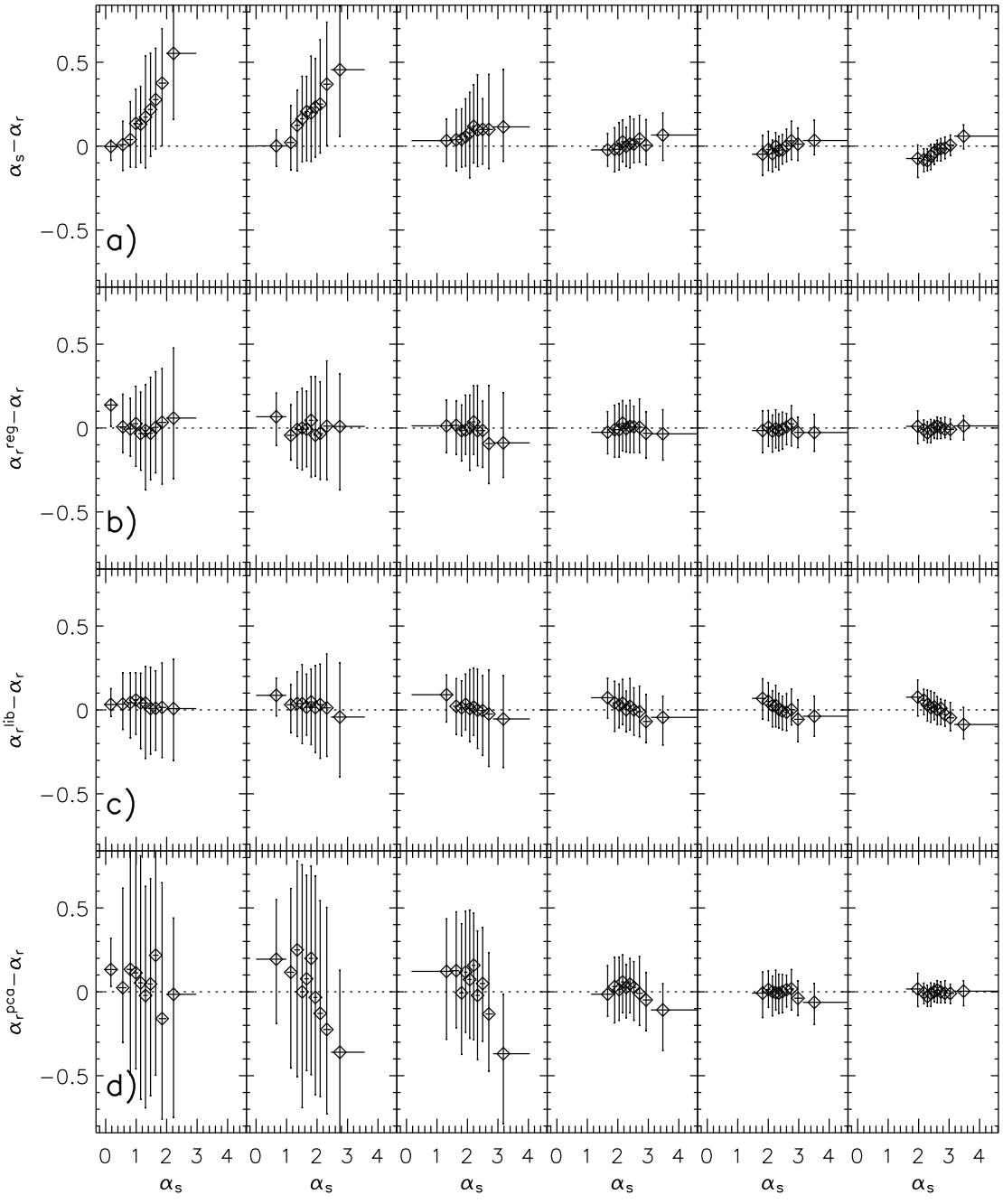

Figure 3.9: The true geometry difference $\Delta_\alpha = \alpha_s - \alpha_r$ as a function of the redshift space geometry α_s (panel a)) for the six different radii (columns). The estimated geometry differences $\Delta_\alpha^{\rm reg} = \alpha_r^{\rm reg} - \alpha_r$, $\Delta_\alpha^{\rm lib} = \alpha_r^{\rm lib} - \alpha_r$, and $\Delta_\alpha^{\rm pca} = \alpha_r^{\rm pca} - \alpha_r$ are presented in the panels b), c), and d), respectively. The data is divided into 10 bins which includes the same number of data points. The abscissa of each data bin is marked with its mean value; the horizontal error bar indicates the bin s distance to the neighbor bin on each side. The ordinate of each bin is marked with its median value; the lower and upper value of the vertical error bars correspond to the 25% and 75% percentile of the data, respectively.

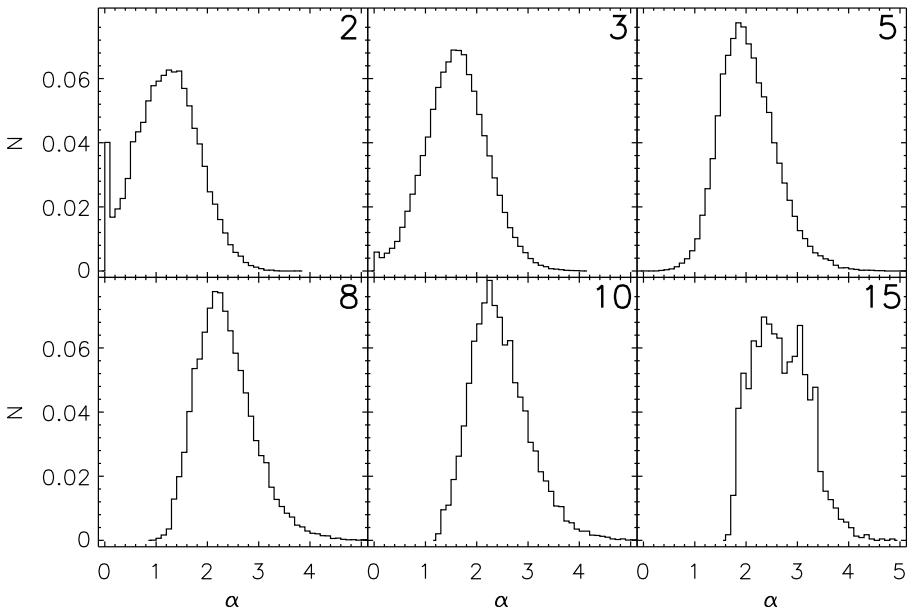

Figure 3.10: Histograms for the geometry estimator α calculated from SDSS galaxy sample for different radii. The radius is given in the upper left corner of each box and bin width is 0 1. Note that the number of galaxies for the different radii is different (see Chapter 2)

homogeneous.

For larger radii $r \geq 10\,h^{-1}$Mpc, the interpretation is different because the low-density galaxies are found in the tail of the distribution. The members of the galaxy clusters are defined through the smaller values because they appear point-like if the sampling volume is larger than the typical cluster size. The intermediate radii $r = 5\,h^{-1}$Mpc and $r = 8\,h^{-1}$Mpc can be regarded as the transition scales between the intra-cluster scales and regions where field galaxies are included in the classification of the environment because of the shape and size of the weighting function.

Although the interpretation for specific values of α can be different for each radius, all α-distributions have in common that their opposite tails are populated with opposite types of galaxy environments, i.e. the high- and low-density regions. The galaxies in between these extreme environments are not only found within regions with a broad range of intermediate density, but also are embedded in environments with distinct geometry properties, e.g. filament-like and wall-like structures, that can be estimated and distinguished with α. The central question of this work, whether this transition of geometry attends with a transition of certain photometric or spectroscopical properties of the galaxy similar to the known correlation between density and physical properties of galaxies, is addressed in Chapter 4.

The calculated α-distributions can be compared with other topology and geometry measures like the Minkowski Functionals that have shown the existence of non-isotropic structure like walls and filaments apart from the isotropic cluster and void-like environments (e.g. Schmalzing & Diaferio 2000). The presented geometry estimator α is in agreement with these results that a spectrum of geometries is found in different galaxy environments, like filaments and sheets. The result from Schmalzing & Diaferio (2000), that most galaxies reside in sheet-like structures (e.g. pancakes), is corroborated by the α-distribution for the radii $r \geq 5\,h^{-1}$Mpc , where galaxies with $\alpha \approx 2$ is the most abundant group. The following section provides an illustration of geometry results found in galaxy distributions.

3.4.2 Illustration in 2D

In order to illustrate the relationship between local geometrical properties in the galaxy distribution and the scaling index α, an representative region was reduced to two spatial dimensions and each galaxy was marked with the individual α-value. In Figure 3.11, a thin slice was extracted from the volume limited SDSS galaxy sample and projected to a plane. Each column corresponds to a certain α-range and each galaxy was marked with symbols if its α-value falls into this range. Additional figures of similar type can be found in the Appendix A.3.

This figure visualizes that different galaxy environment have different geometry properties. Galaxies in dense regions have values of $\alpha \approx 2$ (column C) for $r \leq 3\,h^{-1}$Mpc; this corresponds to an uniform distribution of points in two spatial dimensions. For $r \geq 8\,h^{-1}$Mpc, the identical galaxies have values of $\alpha \approx 0\,5$ (column A), which corresponds to point-like structures. The situation is switched for galaxies in void-like regions: small α-values on small scales (column A), and large α-values on larger scales (column D). The columns in Figure 3.11 thus illustrate the transition between the low- and high-density regions at a specific scale, which was discussed in the previous section. The rows emphasize the transition between the different scales, where the meaning and the interpretation of the geometry estimation changes with the size of the sampling volume.

In between the low- and high-density regions, filament-like structures emerge, which cannot be defined with a certain density level, but which can be classified with the geometry estimator α. For instance, at the radius $r = 8\,h^{-1}$Mpc, the dense cluster regions in column A are connected with filament-like structures of galaxies in column B. Furthermore, galaxies in the filament-like structure surround the galaxies in the void-like regions (columns C and D). Because void and cluster galaxies have very different physical properties, it is interesting to investigate the physical properties of galaxies in filament-like structures, as mentioned in the previous section. Besides, one can investigate the influence of those filament structures on current galaxy formation and evolution models.

Note, that this illustration was performed in two spatial dimensions to simplify the visualization and interpretation. With this example, wall- of sheet-like structures are not present and projection effects might affect the observed structure. The following section will address this point by using dark matter simulations.

3.4.3 Comparison with a Dark Matter Simulation

For a short qualitative legitimation for using a two dimensional (2D) sample for our illustration in the previous section and in order to illustrate some general aspects, the observed galaxy distribution is compared with a simulated dark matter distribution (Huber 2002). From this simulation data, a two dimensional, projected subsample of dark matter points was created which has similar size and number density like the galaxy region used in Section 3.4.2.

In the upper panel of Figure 3.12, the relation between the α-values and the local density contrast δ_{15} calculated for the two dimensional galaxy sample are shown for $r = 15\,h^{-1}$Mpc. Here, $\alpha < 1$ belong to galaxies in clusters, $\alpha \approx 1$ to filament-like structures , $\alpha \approx 2$ to uniform, and $\alpha > 2$ to void regions. The relation between α and the density contrast δ_{15} is not linear. At a constant density level, various geometry values are found that correspond to differently shaped local environments. For instance, similar density levels can be found in clusters, filament-like or uniform regions. This implies that the classification of environments with local density can be degenerated and differently shaped structures in the galaxy distribution cannot be distinguished. With the additional use of the geometry estimator α, complementary information is added, the degeneration is broken, and distinct galaxy environments can be identified. This idea of a more sophisticated description of the galaxies environments is used in the following chapters.

In middle panel of Figure 3.12, the corresponding relation between α and δ_{15} is plotted for a thin slice through the simulated dark matter distribution. Qualitatively similar relations are found in the dark matter distribution and in the observed galaxies distribution. This comparison

=

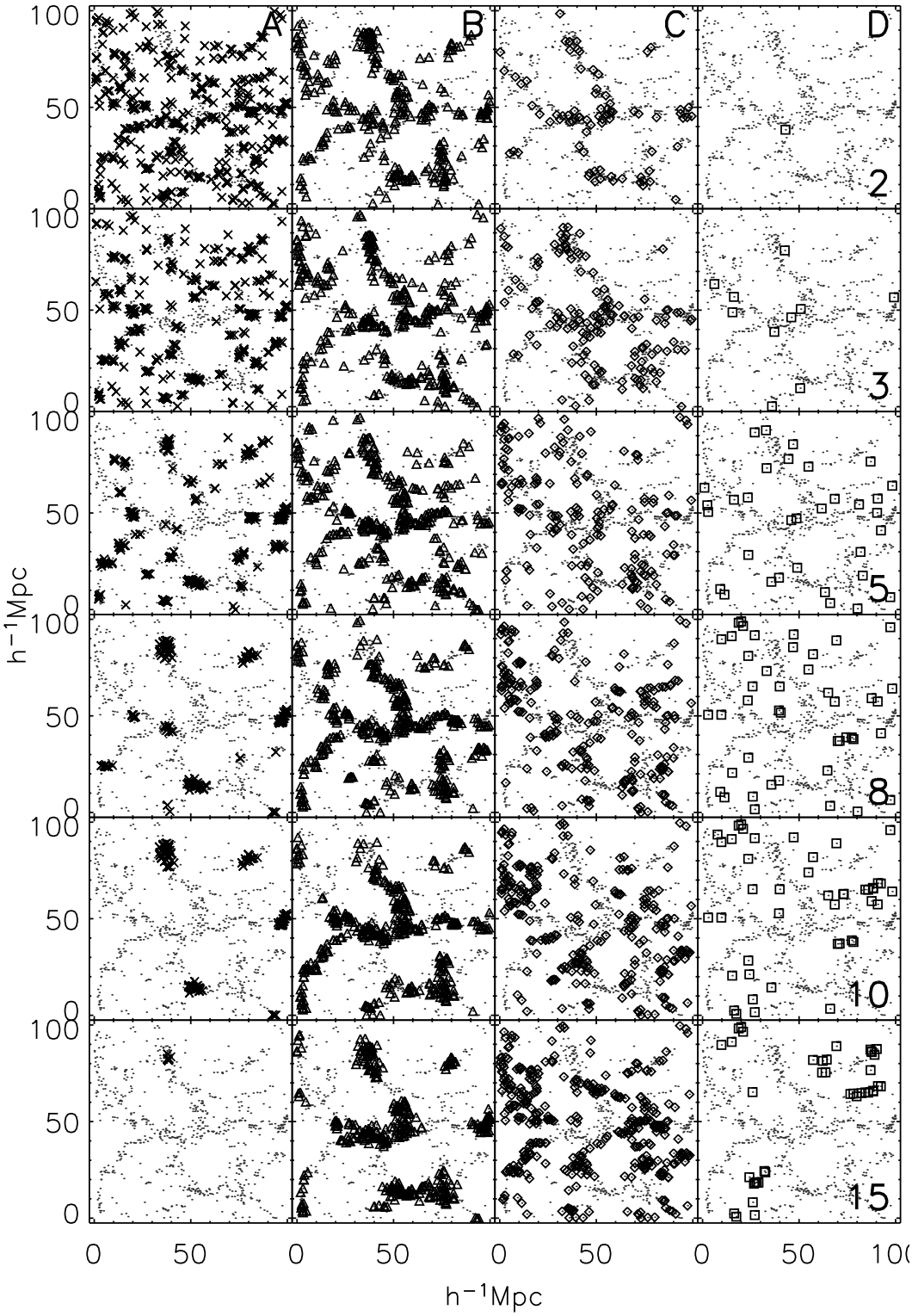

Figure 3.11: Galaxy distribution encoded with different α ranges. A slice of thickness $4\,h^{-1}$Mpc was extracted from the volume limited SDSS galaxy sample, projected to a plane and two dimensional α-values were calculated. The rows correspond to different radii, which are given in the left column. The columns and the symbols correspond to different α-intervals: A: $\alpha < 1\,2$, B: $1\,2 \leq \alpha < 1\,7$, C: $1\,7 \leq \alpha < 2\,4$, and D: $\alpha \geq 2\,4$.

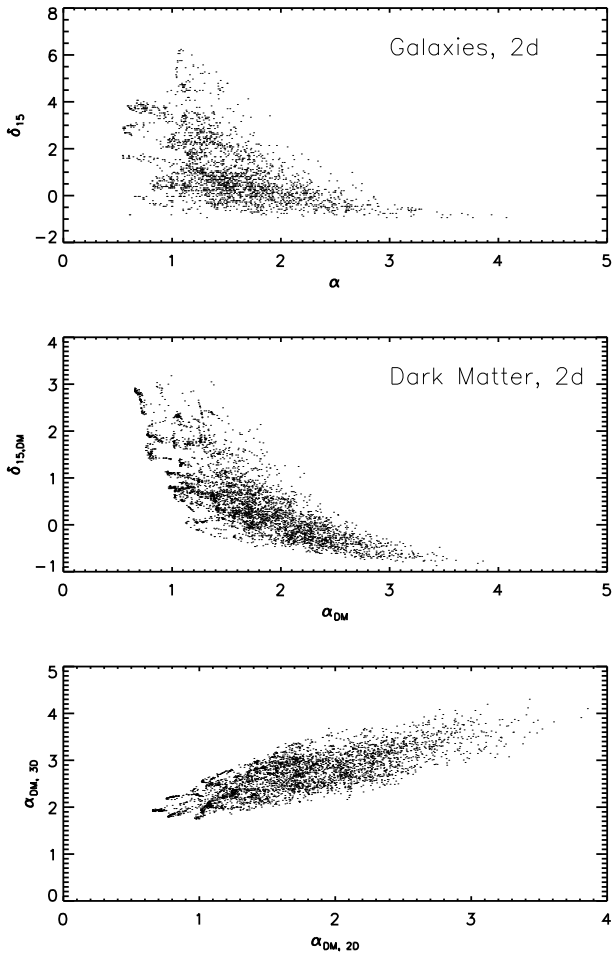

Figure 3.12: Comparison between the geometry-density relation ranges for the observed galaxy distribution (upper panel) and a simulated dark matter distribution (middle panel) in two dimenstion. Lower panel: Comparison between the 2D-α and 3D-α in the dark matter distribution.

suggest that galaxies trace the large-scale structure of the dark matter distribution, and that dark and baryonic matter are found in similar environments. Models of galaxy formation predict that galaxies are created within the density peaks of the dark matter distribution. This bias explains that more low-density regions (voids) are found in the galaxy distribution compared to the dark matter distribution, where even low-density regions are homogeneously populated with dark matter.

The lower panel shows a relation between two-dimensional α_{2D} and the corresponding three-dimensional α_{3D}. Both are strongly correlated and the main difference between 2D and 3D is a shift to higher α-values, $\alpha_{3D} \approx \alpha_{2D} + 1$. The additional dimension in the configuration space increases the geometry values as well, because a random and uniform 3D region corresponds to $\alpha \approx 3$. Because of this correlation, it is legitimate to use a two dimensional analysis to illustrate the meaning of geometry and to gain some useful insights into the description of environments.

Still, the full three dimensional information gained from the galaxy distribution is crucial in order to estimate geometrical features of the large-scale structure. Although low-dimensional surveys may give a good impression about the dominance of the voids and clusters, which are present in e.g. pencil beams as well, the full volume of the SDSS survey is required. Instead of generating a general picture of the texture of large-scale structure, this work intends to classify the environments of galaxies without any projection effects, which can change the structure properties, e.g. introducing more clustering (Kaiser & Peacock 1991; Peacock 1999, Section 16.6).

Regarding the interactions between different types of matter, semi-analytic simulations of galaxy formation indeed predict a direct connection between the galaxy distribution and the underlying dark matter halos. Here, the question emerges at what scales the sampling volume traces the dark matter halos and on what scale one really samples the galaxy distribution. On the small scales, the density or geometry measures estimate an environment that is dominated by galaxies from the same dark matter halo. Galaxies that originate from the same halo presumably share the similar evolution and formation history, and thus have similar physical properties, like the morphology of the galaxy or its star formation history. Therefore, there is rather a strong correlation between the properties of the host dark matter halo (e.g. mass) and its inhabited galaxies than an dependency on galaxy interactions. This relation is observed by estimating the small-scale environment; the effect of the large-scale structure can be excluded. However, with increasing scales, the galaxies from neighboring halos and the neighboring large-scale structure are included in the sampling volume of the estimator. Then, the correlation between galaxy properties and the environment cannot only be explained by the characteristic of the host halo, but interactions between the galaxies have to be considered.

By distinguishing between the influence of the dark matter halo on the small-scale statistic and the influence of the large-scale galaxy distribution on large-scale clustering, the underlying physical processes can also be separated. In the dark matter halos, the effects should be drastic, efficient and fast, like the ram pressure stripping. The classical morphology-density relation is then just a consequence of this model. On the contrary, interactions between galaxies are ineffective and take more time, e.g. tidal interactions.

A even more general aspect from this discussion is the nature or nurture problem in galaxy formation, if galaxies are born in different environments or if they had the identical initial conditions but different evolution histories. By using a small sampling volume, physical properties of galaxies with independent initial conditions (nature) can be investigated, because the formation conditions are defined by dark matter halos. A larger sampling volume includes more effects of galaxy interactions in the past and is thus more sensitive to the condition during the galaxy s evolution (nurture).

3.5 Error Analysis

After the definition of the geometry measure α, it is important to understand the magnitude, dependencies and effects of possible statistical and intrinsic errors that are included its estimation.

The effect of the redshift space distortions, and the stability and the classification error of the geometry estimator are estimated and discussed using SDSS data and simulated galaxy distributions. Edge effects as a possible source of error were excluded by using the minus estimator for the later analysis, i.e. we use only galaxies that have large enough distance to the survey boundary (see Chapter 2).

3.5.1 Stability Tests with Bootstrapping

After the introduction of the scaling index, the statistical errors of this geometry estimator, its stability, are calculated on each scale. The random error is determined to judge how stable the geometry can be measured for a specific galaxy and how sensitive it is towards fluctuations in the environment. For this purpose, the bootstrap method is used, which was first introduced by Efron (1979). More applications in field of astrophysics can be found e.g. by Barrow, Bhavsar & Sonoda (1984), who use the method to analyze the galaxy clustering. An advantage of this approach is that the bootstrap method estimates the errors without any assumptions about the underlying data distribution. It can be used for a broad range of problems, where the standard error cannot be acquired with ordinary methods, e.g. from a set of different realizations or ensembles. A comprehensive description and a comparison with similar methods such as Jackknife can be found in Efron & Tibshirani (1993).

The bootstrap method extract errors from a original data set by creating a set of specific data distributions randomly drawn from the original data, so-called Monte-Carlo simulations or bootstrap samples. For the estimation of the geometrical stability, one Monte-Carlo environment is created for a representative galaxy at $\mathbf{x}_i$ by defining a local volume and by drawing galaxies with replacement out of the original sample around the center of the galaxy. The number of all neighbor galaxies within a distance $2r$, N_α, define the number of the galaxies in each Monte-Carlo environment, the bootstrap sample. Note, that a neighbor galaxy can appear several times in the bootstrap sample, or cannot be represented at all. For each bootstrap sample $b = 1 \quad N_\mathrm{B}$, a geometry value $\alpha_\mathrm{B}(\mathbf{x}_i\ b)$ is calculated for the center galaxy. After repeating the procedure for each of the N_B bootstrap samples, the estimated geometry variance for center galaxy at $\mathbf{x}_i$ can be written as

$$\sigma^2_\alpha(\mathbf{x}_i) = \frac{1}{N_\mathrm{B}-1}\sum_{b=1}^{N_\mathrm{B}}[\alpha_\mathrm{B}(\mathbf{x}_i\ b) - \ \alpha_\mathrm{B}(\mathbf{x}_i)\ ]^2 \tag{3.53}$$

where the mean geometry of the bootstrap samples is given by

$$\alpha_\mathrm{B}(\mathbf{x}_i)\ = \frac{1}{N_\mathrm{B}}\sum_{b=1}^{N_\mathrm{B}}\alpha_\mathrm{B}(\mathbf{x}_i\ b) \tag{3.54}$$

The bootstrap error evaluates how much weight one neighbor galaxy contributes to the total calculation of the geometry value for the center galaxy. With the resampling, some neighbor galaxies are missing or presented several times in some of the bootstrap samples. The corresponding geometry value reflects if and how the geometry depends on single neighbor galaxies. The whole set of bootstrap samples provide an measure of stability and can be used to estimate the variance. One would already expect, that the estimation of α is more stable on larger scales because more neighbor galaxies are included in the sampling volume and results are less dependent on single neighbor galaxies.

In Figure 3.13, the relations between α and the bootstrap error σ_α, and between α and the number of neighbor galaxies N_α (within a distance of $2r$) are plotted. The bootstrap method was performed with $N_\mathrm{B} = 100$ bootstrap samples for each galaxy. In Table 3.6, the statistics of these two relations are summarized. In addition, N_0 denotes the number of galaxies, where too few galaxies, $N_\alpha < 3$, were found within the distance to create bootstrap samples. As expected, the bootstrap error decreases with increasing radius, due to the larger sampling volume and the increasing number of neighbor galaxies. Furthermore, the error is dependent on the environment:

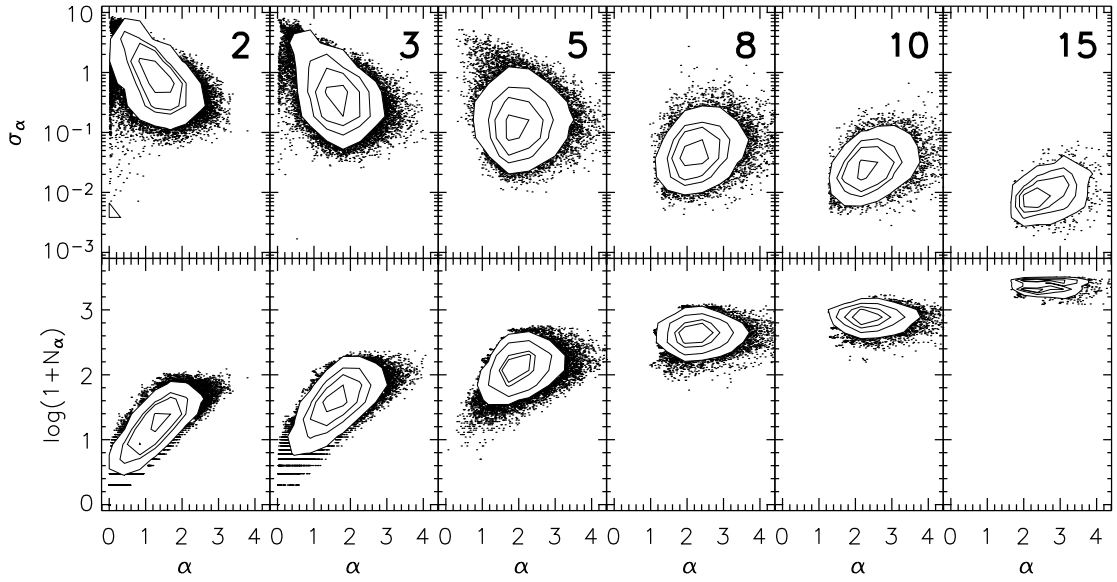

Figure 3.13: The relations between the α and the bootstrap error σ_α (upper panel), and between α and the number of neighbor galaxies N_α (lower panel) for different radii (columns). The four contour lines mark the levels where 90% (outer contour), 70%, 50%, and 30% (inner contour) of all points are included.

within dense cluster regions, the geometry estimation is more stable than for galaxies within void regions which seem to possess a rather fragile geometry.

However, this is a slight drawback of the bootstrap approach that one cannot measure an reliable error for galaxies in void-like environments on the small scales. In these regions, no or only a few galaxy neighbors are found within the chosen distance $2r$ by construction. Either this approach cannot be used to calculate an error, which is the case for some galaxies N_0 (Table 3.6), or large magnitudes of errors were assessed due to a small number statistic and a high sensitivity to outliers. For the N_0 void galaxies an exception rule is introduced: if there are not enough neighbor galaxies within the sampling volume for certain radius, the calculated bootstrap error of the next larger radius is assigned. Because more galaxies are then used at the chosen radius, the resulting error may be underestimated there.

On the other hand, it seems reasonable to assume that a galaxy without any close neighbors has a quite obvious and trivial point-like geometry because the surrounding emptiness of such a galaxy environment is the very definition of a void-like region. Yet, it is necessary to evaluate the information content of each data point in an observed galaxy sample because the observation can produce spurious effects itself including non-existing data points or wrong locations (e.g. redshift space distortions). It is likely, that the presented prescription overestimates the errors in low-density regions and can be regarded as an conservative upper limit. In all the other environments, the bootstrap error provides a good estimate of the local stability which cannot be provided with standard techniques.

Another insight that is gained from Figure 3.13 is the moderate correlation between the geometry α and N_α on small scales ($r \leq 5\,h^{-1}$Mpc). Here, the lowest and highest values, $\alpha < 1$ and $\alpha > 2\,5$ have only a smaller variation within the number of neighbor galaxies. This implies, that the geometry is correlated with the local number of galaxies within a certain distance and, thus, with the local density. However, on the larger scales ($r \geq 8\,h^{-1}$Mpc), the environments with a similar number of neighbor galaxies, were assigned with different α-values and, thus, are not identical and can be differentiated with the geometry information. This result, that the geometry

radius (h^{-1} Mpc)	σ_α	$\log(1+N_\alpha)$	N_0
2	1.102 ± 1.114	0.174 ± 0.375	1983
3	0.650 ± 0.779	0.561 ± 0.340	135
5	0.227 ± 0.300	1.088 ± 0.261	-
8	0.063 ± 0.066	1.618 ± 0.184	-
10	0.034 ± 0.029	1.870 ± 0.137	-
15	0.011 ± 0.008	2.361 ± 0.086	-

Table 3.6: Mean bootstrap errors $\langle\sigma_\alpha\rangle$ for each radius r calculated with 100 Monte-Carlo environments for each galaxy. N_α are the number of neighbor galaxies within the distance $2r$. N_0 denote the number of galaxies with too few neighbor galaxies, $N_\alpha < 3$, in order to calculate the σ_α (see text).

is complementary to the local density and that a full description of environments should include both measures, supports the findings in Section 3.4.3, where a similar conclusion was found.

3.5.2 Classification Errors

After the investigation of the stability, classification rates of the geometry estimator α are quantified based on representative case studies for each environment. The questions arise, how reliable a galaxy s environment can be characterized and what is the rate of the misclassifications. For this purpose, an independent reference classification is required in order to compare it with the results of the geometry analysis. In case of cluster- and void-like regions, it is straightforward because those environment well-defined by their local number density and those regions can be easily distinguished from all other environments. For this reason, those possible environments the local density will be our reference. The situation is different for filament-like and wall-like structures which can only be manually confirmed by an observer.

In this work, a case study is performed to approach this question. Prototypes of the main environment classes defined by their local density are created as a reference. The geometry results from observed environments are then compared with with the results from those reference models. First, the environments are defined by their density and the geometry is calculated. Then, geometry thresholds are defined in order to judge the misclassification on the basis of the density. For this analysis, the simulated galaxy data from the mock catalogs (Section 3.3.2) was used because the classification errors of environments should be estimated independently from other intrinsic errors like the redshift space distortions.

Galaxies within cluster and void-like region were defined by a certain density threshold and then corresponding geometry distributions were calculated. In Figure 3.14, the α-histograms for three environment groups are shown for the different radii. The environments were defined by the following prescription: a cluster galaxy must have a large density contrast at small scales density, $\delta(r = 3\,h^{-1}\text{Mpc}) \geq 10$, whilst a galaxy found in a void-like region has a low density contrast, $\delta(r = 5\,h^{-1}\text{Mpc}) \leq -0\,2$. Those thresholds correspond with the values obtained by Rojas et al. (2005). All other galaxies are embedded in presumably filament-like, wall-like, and uniformly distributed regions. The geometry distribution of the cluster galaxies (solid lines) represent the largest values on the smallest scale, includes intermediate values at the transition radius $r = 5\,h^{-1}\text{Mpc}$, and mainly consist of smaller values for radii $r \geq 8\,h^{-1}\text{Mpc}$. This shift of the cluster histogram again reflects the insight, that interpretation of the geometry is scale dependent. On small scales, the galaxies inside of clusters are almost homogeneously distributed. From outside, most cluster appear not perfectly spherical, but more like flattened ellipsoid which resembles a thick disc. This explains why the mean geometry value of cluster is $\alpha(r = 2\,h^{-1}\text{Mpc}) \approx 2$. On the larger scales $r \geq 8\,h^{-1}\text{Mpc}$, the point-like concentrated number density of galaxies lead to

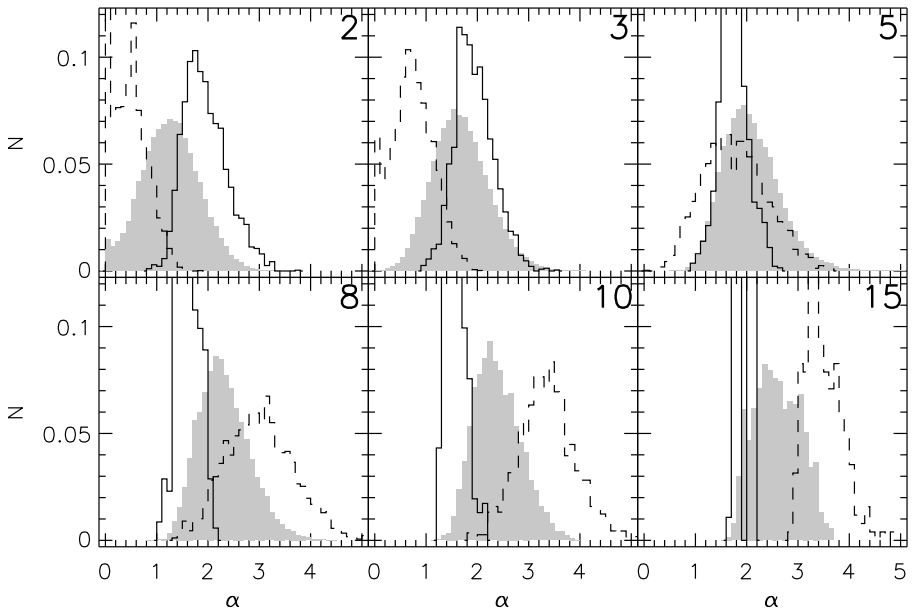

Figure 3.14: The α-distributions for cluster galaxies (solid lines), galaxies in void-like regions (dashed lines), and galaxies in the remaining environments (shaded area) for different radii.

the smallest geometry values within the whole distribution. The mean values of $\alpha \approx 1\,5$ can be explained by the fact that the clusters are embedded in the cosmic web of filaments and walls, which contribute to the geometry estimation on these scales as well.

One advantage of the scaling index method can be that it assigns similar values to galaxy groups and clusters, although both have different local densities. If a galaxy group has a slightly lower density than the chosen threshold, it will not be included in the cluster sample, although the group galaxies might have similar properties compared to the cluster galaxies. A full classification cannot be expected because if one uses density as a reference because there is no clear threshold between cluster and non-cluster galaxies. Either a cluster sample is not complete, or it is contaminated with more galaxy groups.

Galaxies in void-like regions (Figure 3.14, dashed lines) have a point-like geometry due to no or only a few neighbors galaxies at small scales $r \leq 3\,h^{-1}$Mpc . Again, the radius $r = 5\,h^{-1}$Mpc represents a transition scale, where the sampling volume is large enough to include the surrounding galaxies at the boundary of each void. With increasing radius ($r \geq 8\,h^{-1}$Mpc), this effect continues for galaxies in the underdense regions which then have a mean geometry of $\alpha \approx 3\,5$.

The remaining galaxies (gray shaded histogram) are embedded in filament-like, wall-like, and uniformly distributed regions of space. The geometry distribution of those galaxies remains mainly between the two extreme density classes, cluster and void galaxies. Its mean geometry increases form $\alpha \approx 1$ at $r = 2\,h^{-1}$Mpc, to $\alpha \approx 2\,5$ at at $r = 2\,h^{-1}$Mpc.

Regarding the potential of separating the three classes with the geometry α, cluster and void galaxies can easily be distinguished on most scales except for $r = 5\,h^{-1}$Mpc, where all three distributions have a large overlap. The third class, which includes filaments, walls, and uniform regions, can best be discriminated from the other two classes on the smallest scale and for $r \geq 8\,h^{-1}$Mpc. It is interesting that the transition $r = 5\,h^{-1}$Mpc is close to the correlation length r_0 in Equation (3.5) of the two-point galaxy-galaxy correlation function. This length marks the transition between the linear and non-linear regime in the description of the galaxy clustering. It is assumed that the linear regime can be well-described with any second moment statistic. However, the presented histograms imply that even on larger scales, which are considered to be part of

the linear regime, the assumptions of an isotropic environment might not be appropriate. In this regime, the geometry α that includes higher statistical moment is still able to distinguish between different existing environment classes.

3.5.3 Case Studies with Geometry Prototypes

After using the simulated data from the mock catalog, selected geometry prototypes are extracted form the projected, two-dimensional SDSS galaxy distribution. These prototypes are elements of the galaxy distribution that are dominated by one of the typical geometry classes. A set of prototypes was defined and selected by visual inspection for cluster, filament-like, void-like and uniformly distributed regions. Galaxies in wall-like regions were excluded in this two-dimensional illustration.

In Figure 3.15, four cluster prototypes are presented with the surrounding galaxy distribution and corresponding geometry for the radii $r = 2\,h^{-1}$Mpc and $r = 15\,h^{-1}$Mpc. All cluster regions have mostly uniform values for radii smaller than the cluster-size ($2\,h^{-1}$Mpc). The prototypes are embedded into the surrounding cosmic web and are not disconnected, individual parts of the large-scale structure. Hence, the point-like geometries of the cluster itself is embedded in mostly filamentary environment which leads to the observed mode $\alpha \approx 1$ at the larger radius. In addition to this effect, redshift space distortions elongate the cluster which can be seen at the Prototypes 2 and 4. The richness of the Prototypes 3 and 4 is small and those examples represent galaxy groups, which have a similar geometry to clusters but a lower local density.

Prototypes for galaxies in filament-like regions in Figure 3.16 possess string-like values around $\alpha \sim 1$. The chosen prototypes have a clear string-like orientation due to the gravitational effects during the evolution of the large-scale structure. Since most filaments structures connect two distant galaxy clusters, it is not obvious to decide whether a galaxy is still part of the filament or is already located in the outskirts of a cluster. In addition, as mentioned above, clusters and groups of galaxies can be part of filament-like structures. There will be always a fuzzy transition between cluster and filament galaxies close to high-density regions, which complicates a clear classification. However, in the low-density regions, there are still galaxies, which are aligned onto a string-like structure and the geometry can distinguish those galaxies from others, as seen in Prototype 1. At the larger scale, the dominant value is $\alpha \approx 1$. The other prototypes have slightly higher values of $\alpha \approx 1\,5$ due to the fact that the filaments are connected to different environments. Here, the small scale geometry provides no robust information, because the small sampling volume around each filament galaxy does not include the complete structure formed by neighbor galaxies.

Uniformly distributed regions of space can exist on almost all scales in the large-scale structure as well as in the local environment and certainly on the large scales, where the Universe becomes uniform. In Figure 3.17, prototypes of the uniform regions in the local environment of galaxies are shown. In the projection, environments of the black marked galaxies have no well defined geometry. However, at both radii, all prototypes have a similar mean geometry, which indicates a scale invariant geometry that is expected for uniform regions. Prototype 1 and 2 are similar to a broad filaments, while the prototypes 3 and 4 are part of large wall. The full three dimensional information is needed in order to clearly distinguish whether a galaxy is member of a wall, filament-like or a cluster regions.

The prototypes of galaxies in void-like regions in Figure 3.18 reside in clearly low density regions. Large geometry values $\alpha \geq 3$ were calculated at the large scales. At the small radii $r = 2\,h^{-1}$Mpc, a small geometry was found because few neighbors galaxies lead to the point-like geometry. The radius where the transition between this both extremes geometries occurs can be considered as a rough estimate for size of the void region. For instance, Prototype 2 is the smallest void region, which leads to a lower geometry on large scales compared to the others prototypes. Beyond the void region, it is quite common that those void galaxies are surrounded by filament-like structures and clusters.

In Figure 3.19, the α-distributions for each prototype class are summed up from the previous

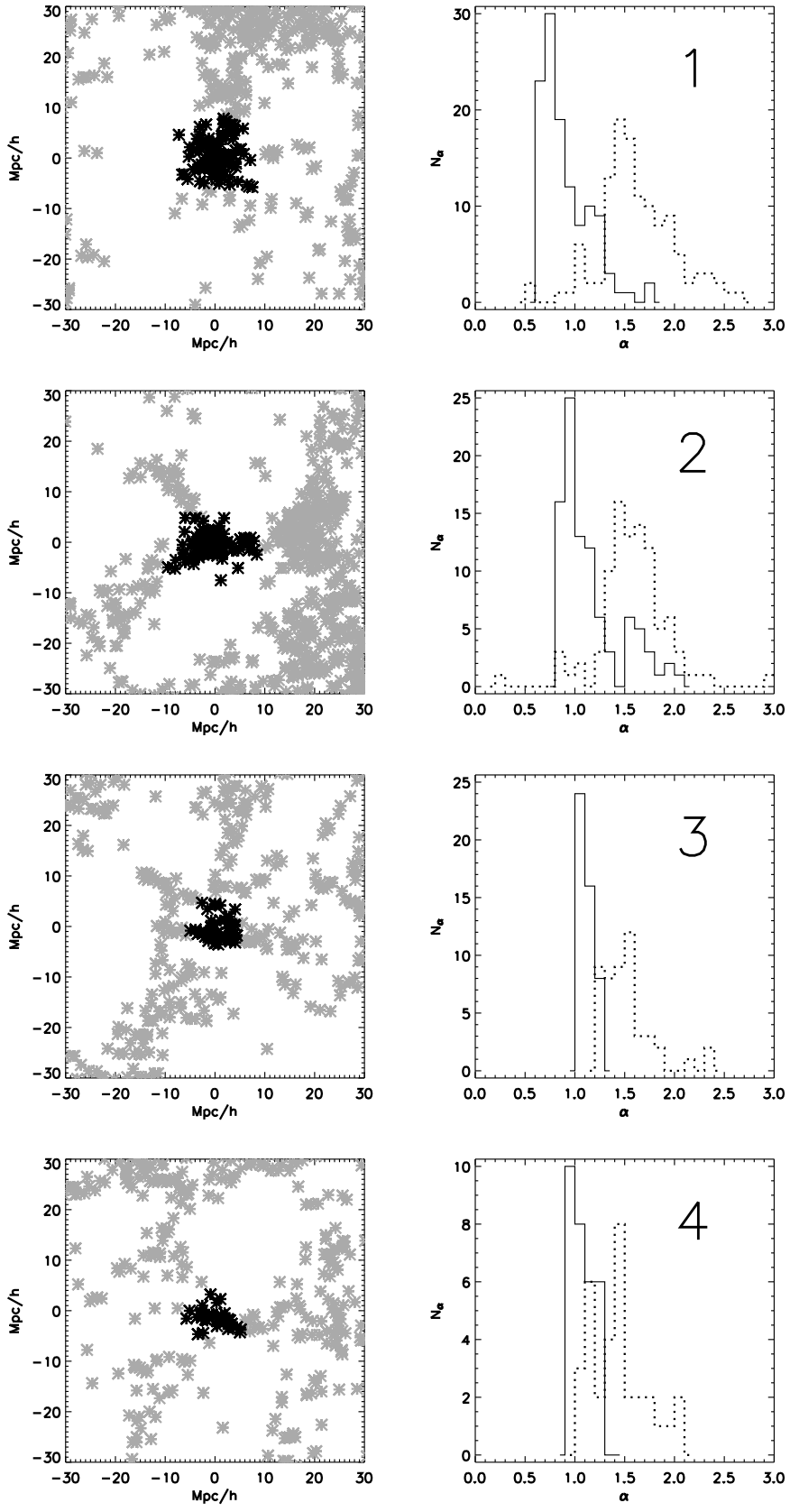

Figure 3.15: Four cluster prototypes (black) with the surrounding galaxy distribution (gray). The right column shows the histograms for the prototype s geometry α for $r = 2\,h^{-1}$Mpc (pointed line) and $r = 15\,h^{-1}$Mpc (solid line).

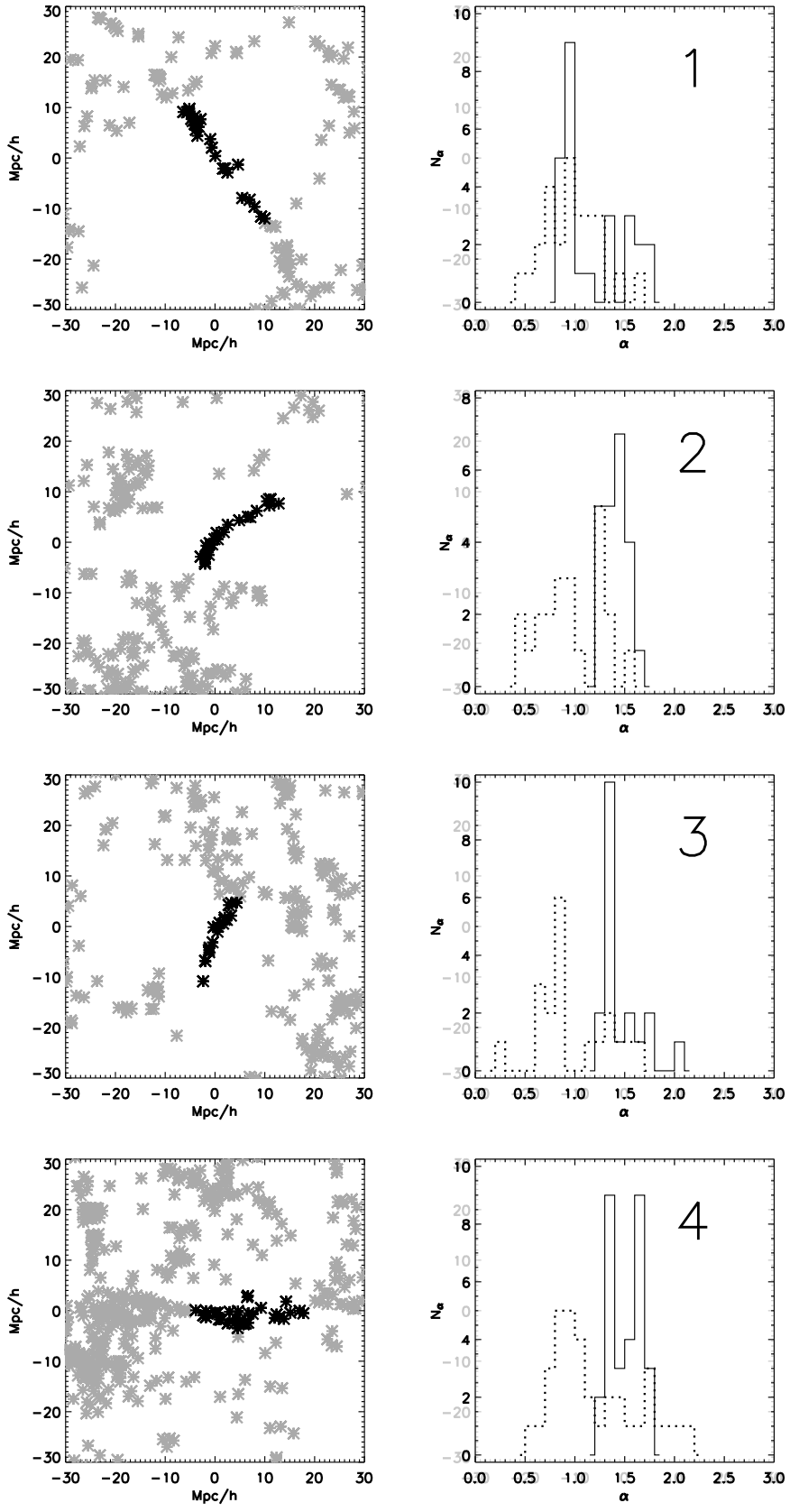

Figure 3.16: Four prototypes for galaxies in filament-like regions (black) with the surrounding galaxy distribution (gray). The right column shows the histograms for the prototype s geometry α for $r = 2\,h^{-1}$Mpc (pointed line) and $r = 15\,h^{-1}$Mpc (solid line).

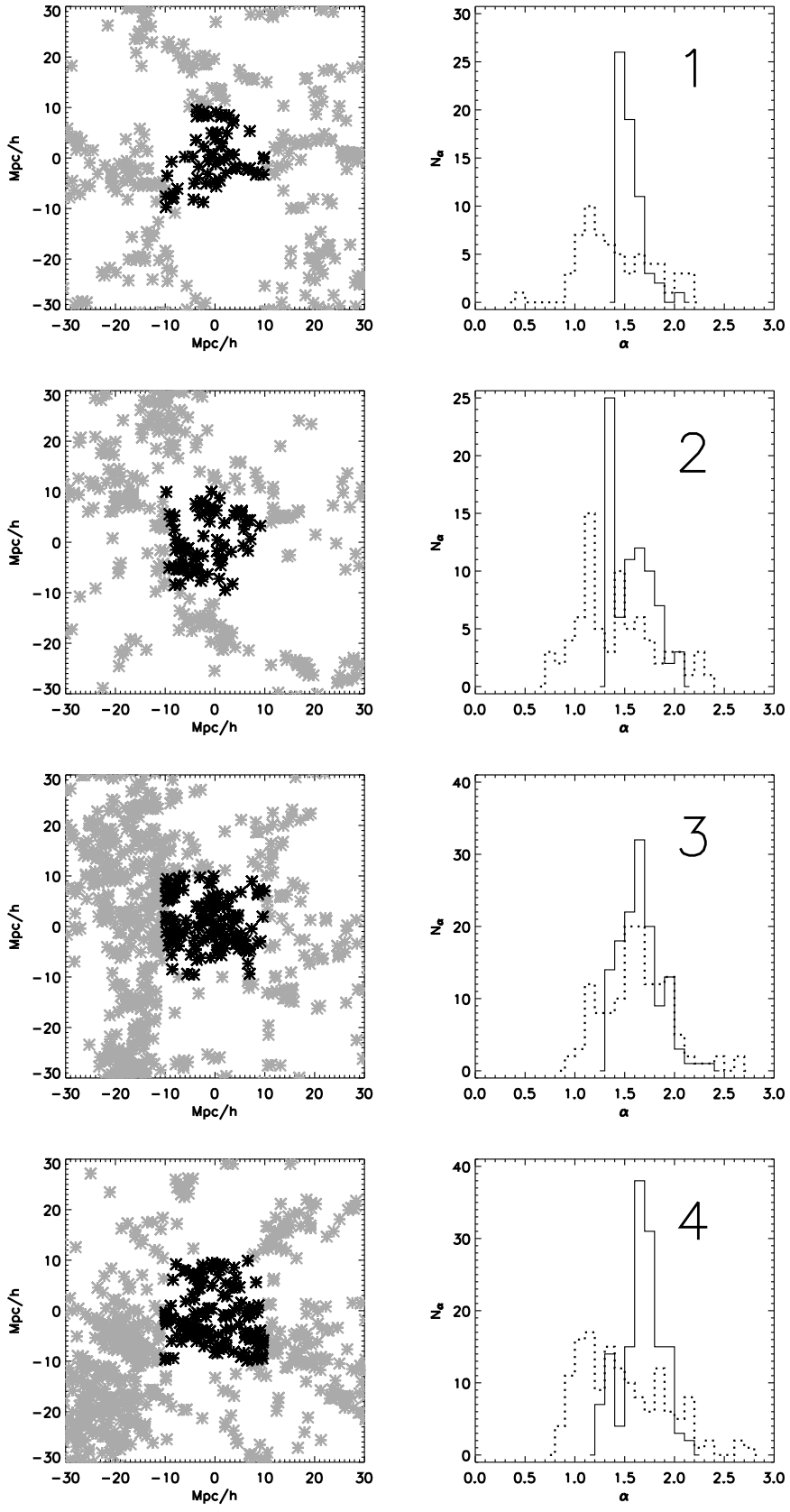

Figure 3.17: Four prototypes for galaxies in uniform-like regions (black) with the surrounding galaxy distribution (gray). The right column shows the histograms for the prototype s geometry α for $r = 2\,h^{-1}$Mpc (pointed line) and $r = 15\,h^{-1}$Mpc (solid line).

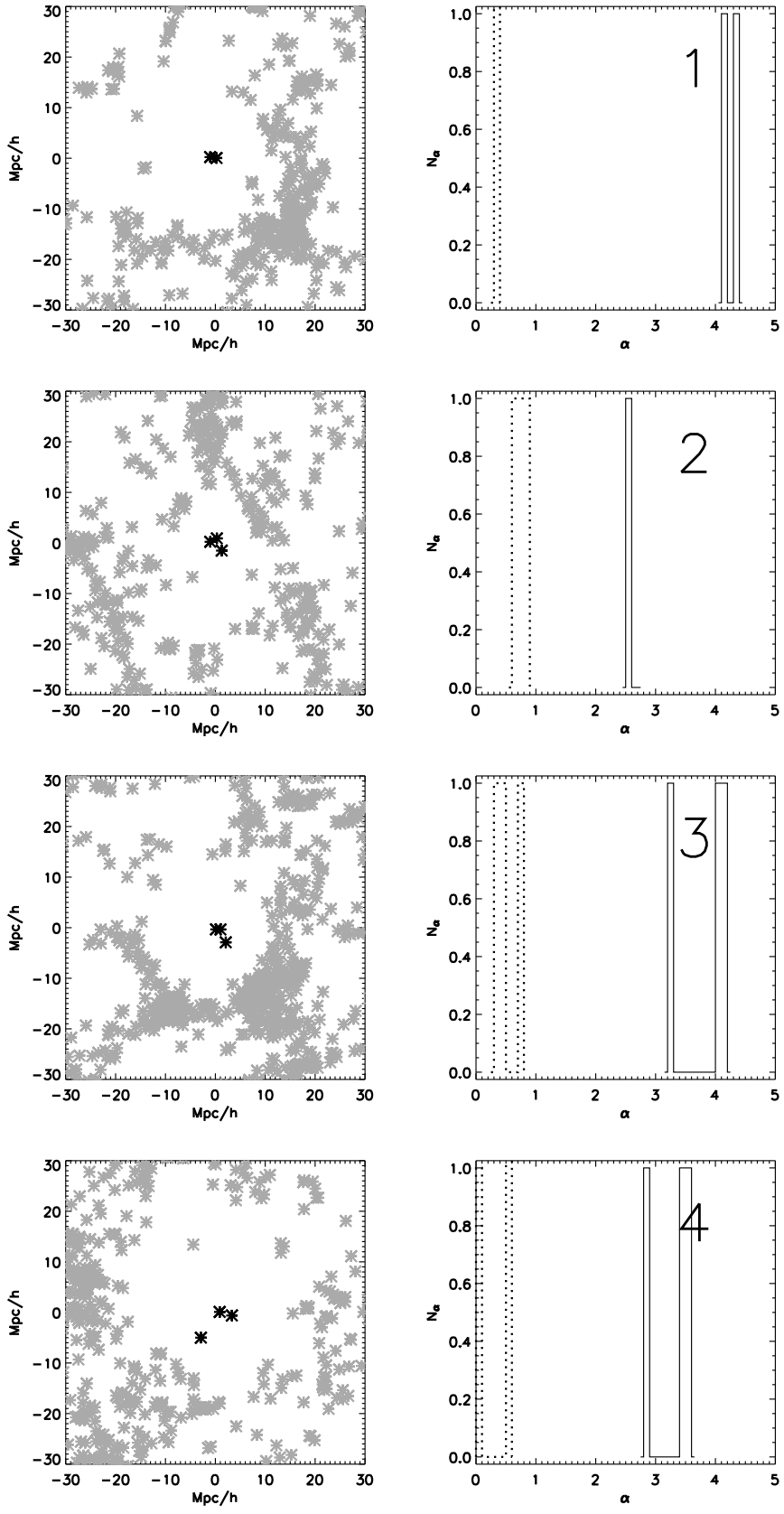

Figure 3.18: Four prototypes for galaxies in void-like regions (black) with the surrounding galaxy distribution (gray). The right column shows the histograms for the prototype s geometry α for $r = 2\,h^{-1}$Mpc (pointed line) and $r = 15\,h^{-1}$Mpc (solid line).

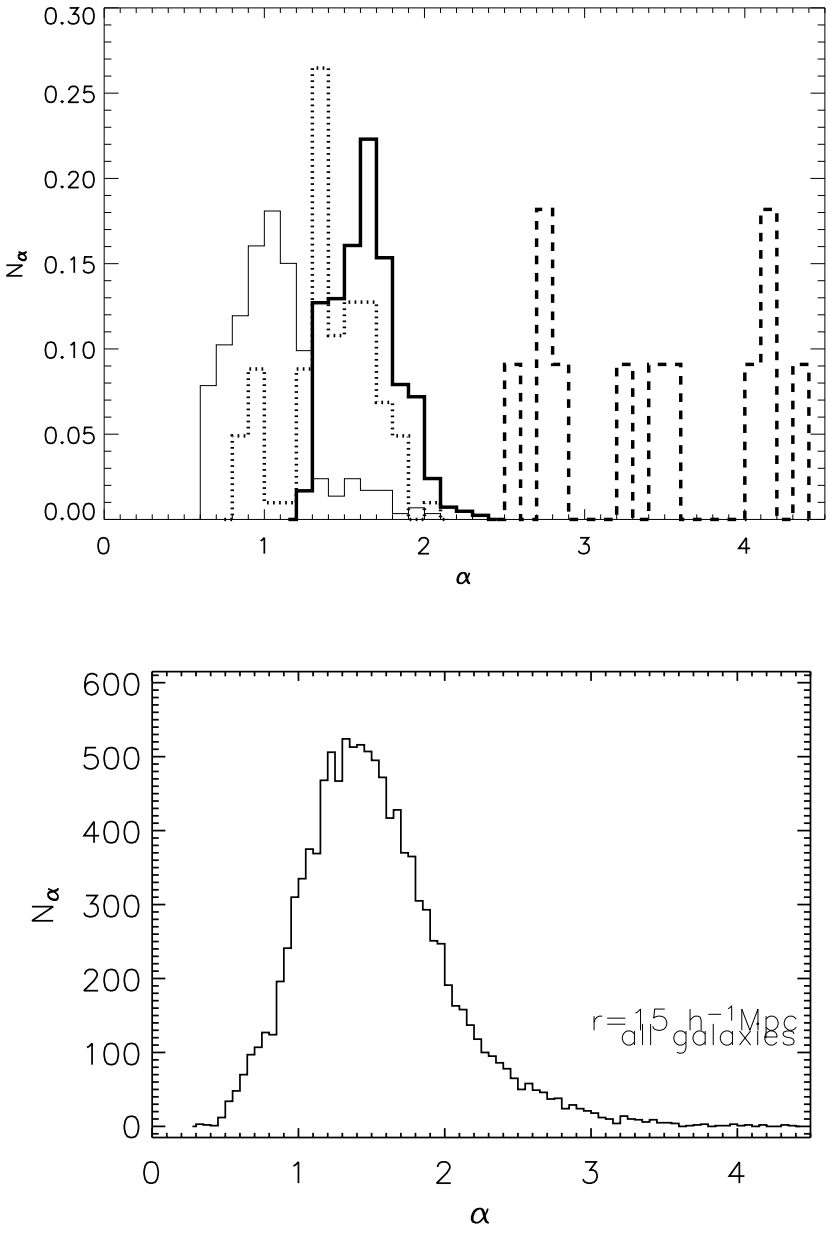

Figure 3.19: Summary of all geometry values of the each of the four prototypes. In the upper panel, each histogram corresponds to one environment: clusters (solid thin line), filaments (dotted line), uniforms (solid thick line) and voids (dashed line). The lower panel shows the α-distribution for the whole sample.

illustrations for the radius $r = 15\,h^{-1}$Mpc . The geometry distributions easily distinguish between the cluster and void-like regions, as well between those both groups and the rest. The distinction between the filaments and the uniformly distributed environments is less obvious. Due to the projection in these illustrations, the full three dimensional information that is contained in the three-dimensional sample is not available. In addition, the choice of the those environments prototypes was performed manually and might be suffer from selection effects. Although this weakness of the presented illustration, insights into the meaning of the geometry α are gained, which supports the principal existence of differently shaped environments.

A more quantitative reference might be useful and could be used in order to separate one-dimensional filaments, two-dimensional sheets and walls and the three dimensional uniformly distributed regions. Such a reference could be established by a cross validation with other non-linear methods that also estimate the geometrical or topological features of the galaxy distribution, e.g. the Minkowski Functionals. However, most of those methods cannot be applied for a local environment, like the geometry estimator α does, and only calculate global estimates. Thus, a direct comparison of prototypes with other methods is difficult, although the global results are in agreement with them.

Chapter 4

Results

In this chapter, the relations between the environment of galaxies and their photometric and spectroscopic properties are presented. For the used volume-limited SDSS galaxy sample (Section 2.2), the change of the luminosity, color, morphology, and indicators of star formation with the local geometry and the local density is shown.

4.1 Luminosities

The luminosities of the galaxy is one of the basic photometric observable. Divided into the five bands, the absolute magnitudes may provide already a first hint towards environment dependent star formation. It is assumed that e.g. the blue band continuum, M_u, reflects a blue population of star within the galaxies. Those stars were formed recently and indicate an active star formation due to a high abundance of gas.

In Figure 4.1, the absolute magnitudes of the five SDSS bands, M_u, M_g, M_r, M_i, and M_z, are plotted as a function of the local density contrast δ for each radii. Galaxies in dense regions of space have a higher blue band magnitudes M_u, i.e. emit less luminosity, compared to galaxies in low density regions. For the red bands M_r, M_i, and M_z, the opposite is found: galaxies in high-density environments emit more luminosity than galaxies in low-density environment. The intermediate band M_g shows only a slight correlation for the latter trend. Infrared light from the z-band can indicate the presence of dust which absorbs ultraviolet starlight from young massive O- and B-stars and emits this energy in as thermal radiation at infrared wavelength. Since this effect would reduce the correlations, the shown relation are rather a lower limit.

All the observed correlations between the local density and the luminosities are most prominent on the smaller scales $r \leq 5\,h^{-1}$Mpc, where the density is sensitive to the dark matter halos of massive clusters which host luminous red galaxies. On the larger scales, the correlation decrease because the large sampling volume smoothes the local environment, and the scatter of the bins increases due to the lower number of galaxies.

In Figure 4.2, the absolute magnitudes are plotted as a function of the local geometry α. Again, the galaxies in cluster environments ($\alpha > 2$, $r = 2\,h^{-1}$Mpc; and $\alpha < 2$, $r \geq 8\,h^{-1}$Mpc) emit less light in the blue band M_u, and less light in red bands M_r, M_i and M_z, compared to galaxies in void-like regions ($\alpha < 0\,5$, $r = 2\,h^{-1}$Mpc; and $\alpha > 3$, $r \geq 8\,h^{-1}$Mpc). Only a minor relation was found for the M_g. The correlations are slightly weaker for the smaller scales compared to the local density, but slightly stronger on the scales $r \geq 8\,h^{-1}$Mpc. For the relations at $r = 3\,h^{-1}$Mpc and for u-band at $r = 5\,h^{-1}$Mpc, no significant trends were found.

Galaxies with a local density in an intermediate range are found in between the two extreme density regions. The geometry estimator assigns them different values that correspond to galaxies in filament-like, wall-like or uniformly distributed environments. Regarding the luminosity, those galaxies can be defined as a transition class between galaxies in clusters and galaxies in void-like

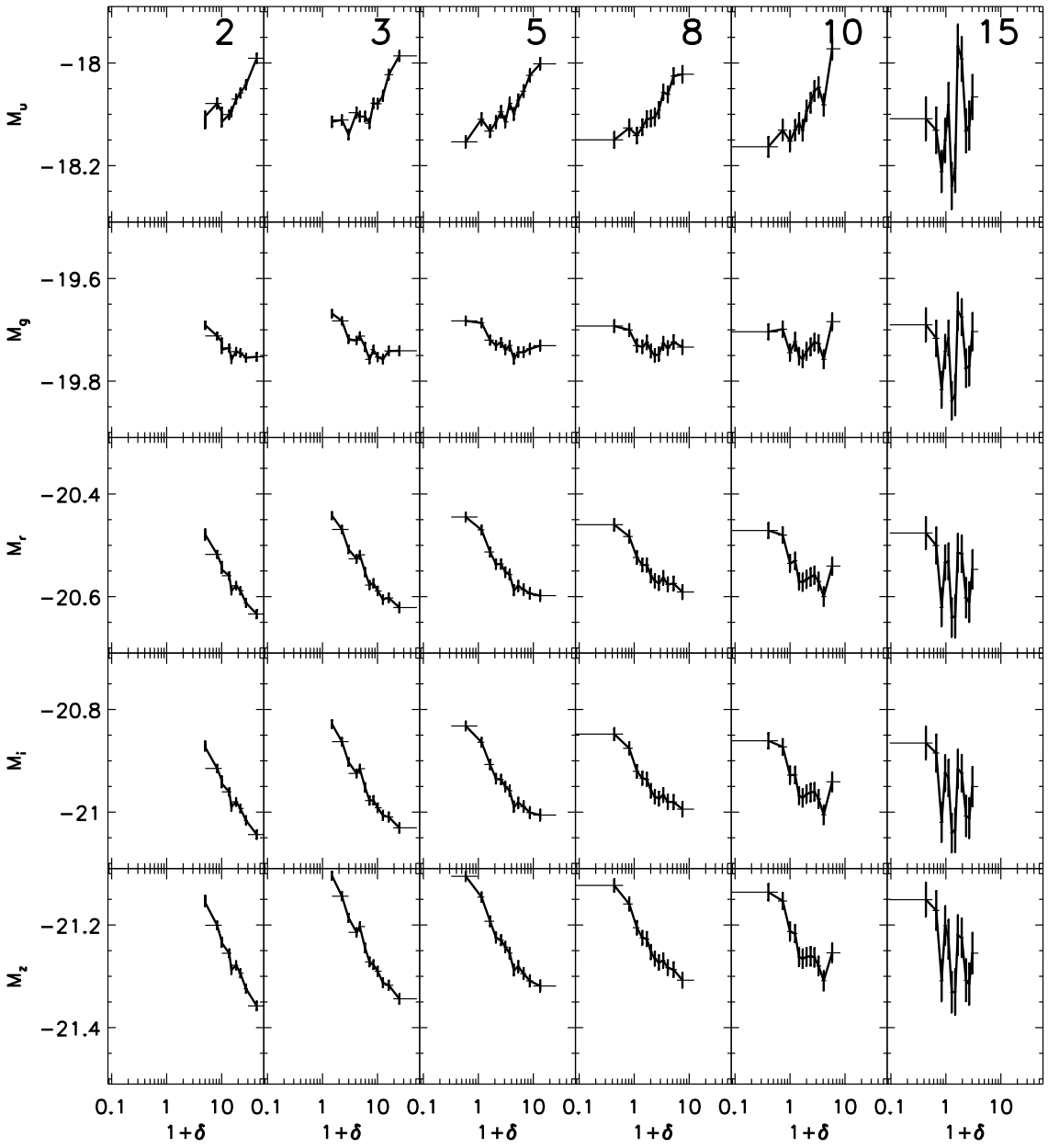

Figure 4.1: The mean absolute magnitudes M_u, M_g, M_r, M_i, and M_z as a functions of the local density contrast δ at the given radii. In each subplot, the data is represented by 12 bins with an equal number of data points. The horizontal error bars of each bin indicate the width; the vertical errors of the y-axis the formal 1σ Poisson error.

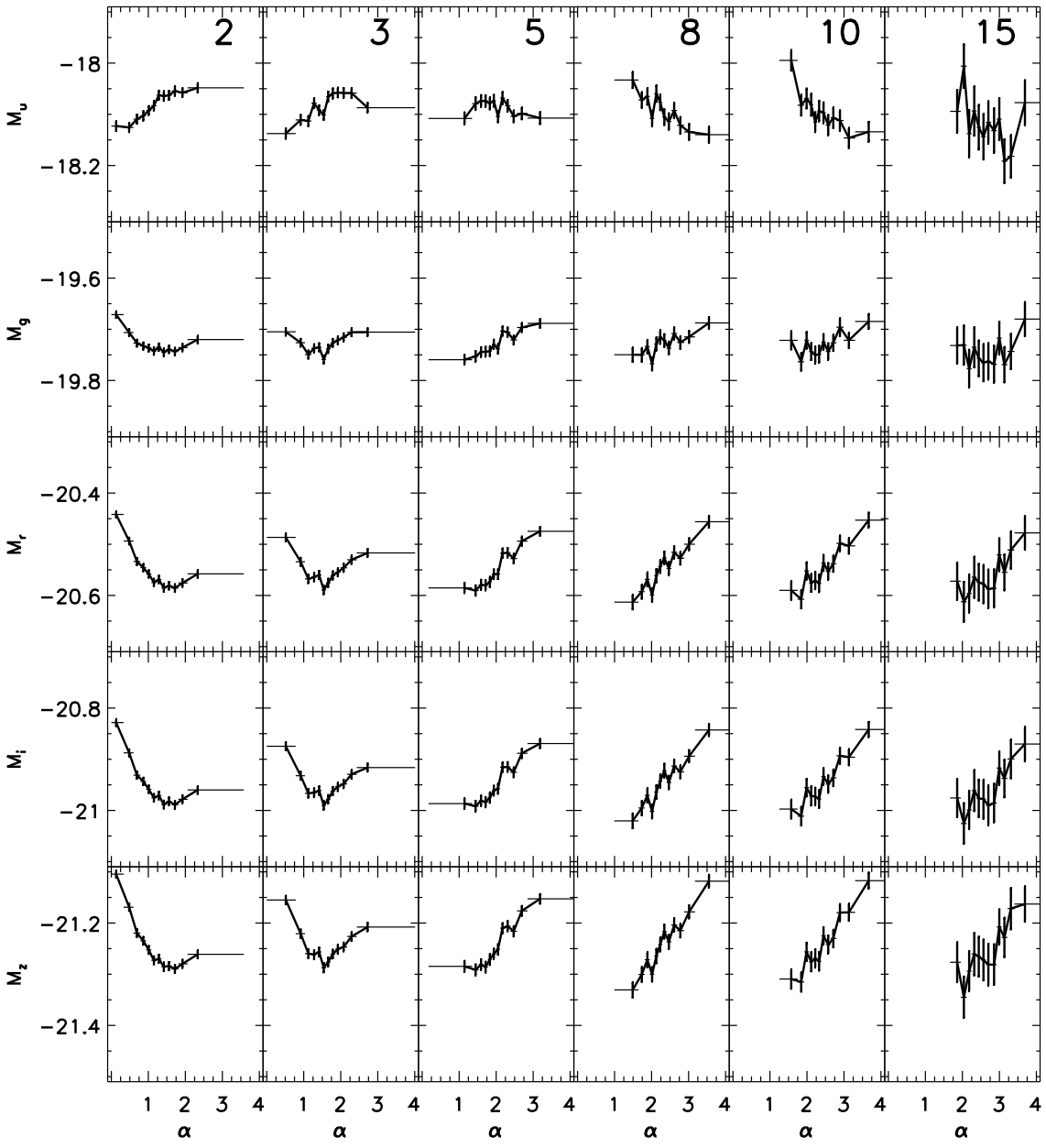

Figure 4.2: The mean absolute magnitudes as a functions of the geometry estimator α at the given radii. In each subplot, the data is represented by 12 bins with an equal number of data points. The horizontal error bars of each bin indicate its width; the vertical errors of the y-axis the formal 1σ Poisson error.

regions. Galaxies in filament-like and wall-like regions are closer to the class of cluster galaxies and show are redder spectrum than galaxies in void-like regions. Galaxies in the uniformly distributed regions are closer to the class of void galaxies and emit more blue band luminosity than the cluster galaxies. This indicates that different parts of the visible large-scale structure emit different light which can be correlated with the local geometry.

4.2 Galaxy colors

In Figure 4.3, the relations between the galaxy colors $u-r$, $g-r$, $r-i$, $r-z$, and $i-z$ and the local density contrast are shown. On all scales, red galaxies that emit more luminosity at shorter wavelength, e.g. $u-r > 2\,4$ and $g-r > 0\,8$, are found in environments with higher densities and blue galaxies are located in low density regions. This is in agreement with the results obtained by Strateva et al. (2001) and Baldry et al. (2004). It also fortifies the results from the luminosity relations and confirms the models of galaxy formations, where galaxies in clusters are dominated by old, red stars and, in contrast, that the galaxies in void-like regions emit more blue light from younger, recently formed stars. Note, that the more red galaxies are found in the galaxy sample. Due to the binning with a constant number of bins, the red color ranges, e.g. $u-r > 2\,5$, are represented with more bins than the galaxies with blue colors.

The relation between the colors of galaxies and the local geometry in shown in Figure 4.4. There is a clear relation between the mean intrinsic color of galaxies and the mean geometry of its environment. All scales indicate that cluster regions mainly consist of red galaxies, void-like regions of blue galaxies, and that filament-, wall-like and uniformly distributed regions represent the transition geometries with a mix of both colors. The color and geometry are stronger related than the luminosity and geometry. In general, the color index is a preferred way to describe the stellar population of galaxies because it is less sensitive to the size and the total luminosity of galaxies.

4.3 Morphology

The concentration index *cin* defined by the bulge-to-disk ratio r_{50} r_{90} (Section 2.6), is used to examine the relation between the estimated galaxy morphology and its environment. The photometric property *cin* indicates the visual shape of galaxies which can be divided in two groups: ellipticals, with a concentrated luminosity profile, have a low value of *cin* and spirals, with a more diffuse shaped profile, have a higher value of *cin*. Note that the concentration index cin has not been corrected for any seeing effects. However, as shown by Goto et al. (2003b), the seeing dependence of *cin* in the SDSS data is weak for galaxies in the redshift interval $0\,05 < z < 0\,1$, which is very close to the redshift range adopted in this work.

4.3.1 Concentration Index

In Figure 4.5, the mean concentration index is plotted against the density contrast for several radii. On all scales, galaxies with a decreasing bulge-to-disk-ratio are found along denser environments. This common trend is consistent with the well-known morphology-density-relation (Dressler 1980; Goto et al. 2003a). The relation is prominent on small scales $r \leq 5\,h^{-1}$Mpc, typical sizes of clusters, and becomes slightly shallower on larger scales, where the larger sampling volume underestimates high-density regions.

The relation between mean geometry α and mean concentration index *cin* is shown in Figure 4.6. The correlation between those two properties is again scale-dependent. On the smallest radius $r = 2\,h^{-1}$Mpc, we see two regimes, one for $\alpha < 1$, where *cin* decreases with growing α, and one for $\alpha \geq 1$, where *cin* stays constant. Mainly spiral galaxies with larger values of *cin* are found below $\alpha < 1$, thus their environment is mostly empty within this small sampling volume. These galaxies

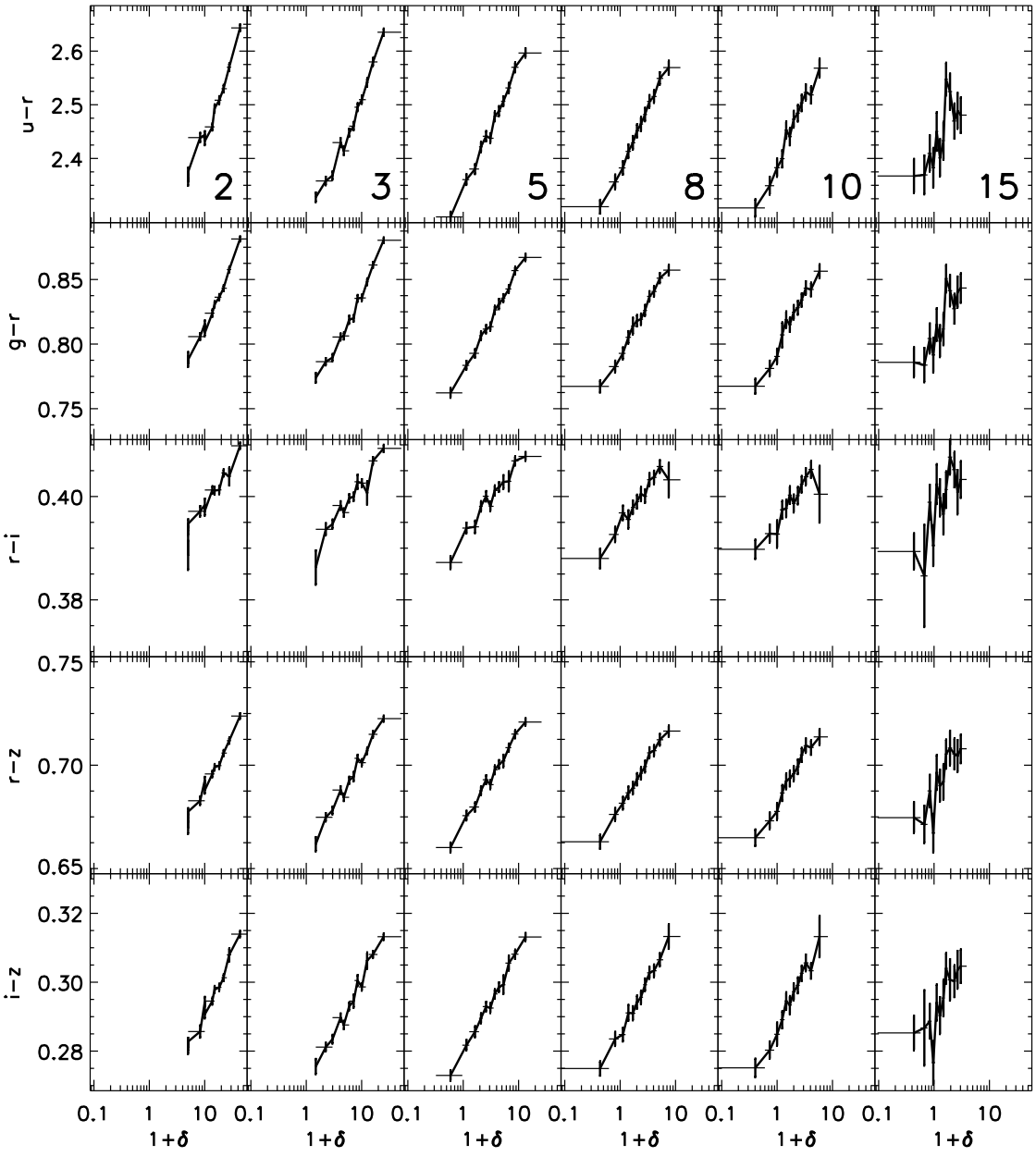

Figure 4.3: The mean colors index as a functions of the mean galaxy density contrast δ at the given radii. In each subplot, the data is represented by 12 bins with an equal number of data points. The horizontal error bars of each bin indicate its width; the vertical errors of the y-axis the formal 1σ Poisson error.

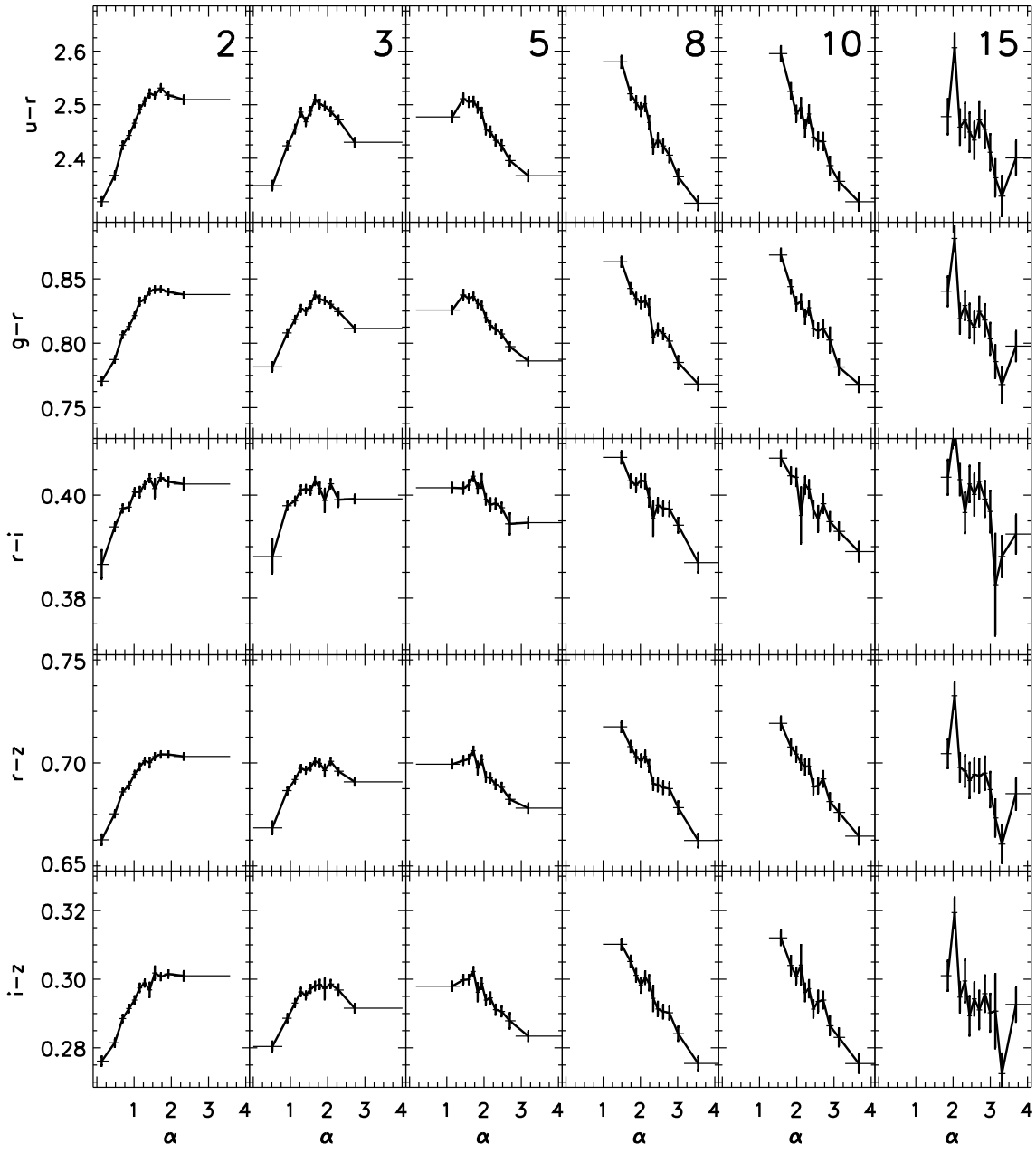

Figure 4.4: The mean colors index as a functions of the geometry estimator α at the given radii. In each subplot, the data is represented by 12 bins with an equal number of data points. The horizontal error bars of each bin indicate its width; the vertical errors of the y-axis the formal 1σ Poisson error.

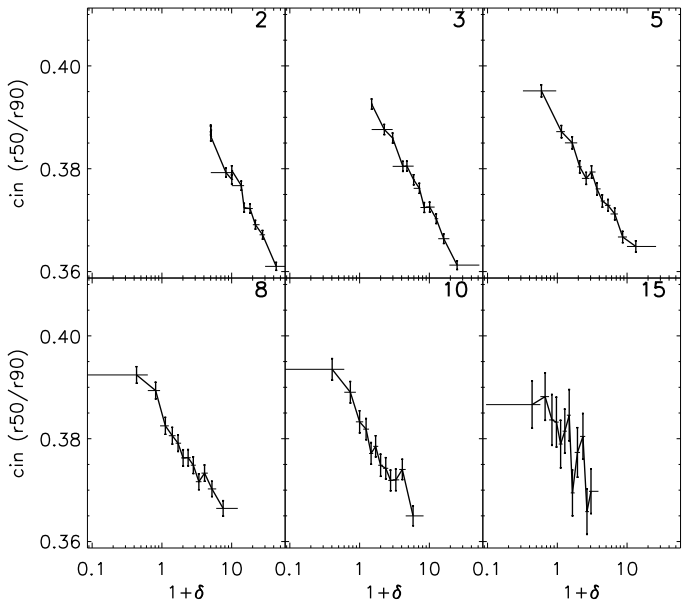

Figure 4.5: The mean concentration index *cin* as a functions of the mean galaxy density contrast δ at the given radii (upper left corner of each panel). In each subplot, the data is represented by 12 bins with an equal number of data points. The horizontal error bars of each bin indicate its width; the vertical errors of the y-axis the formal 1σ Poisson error.

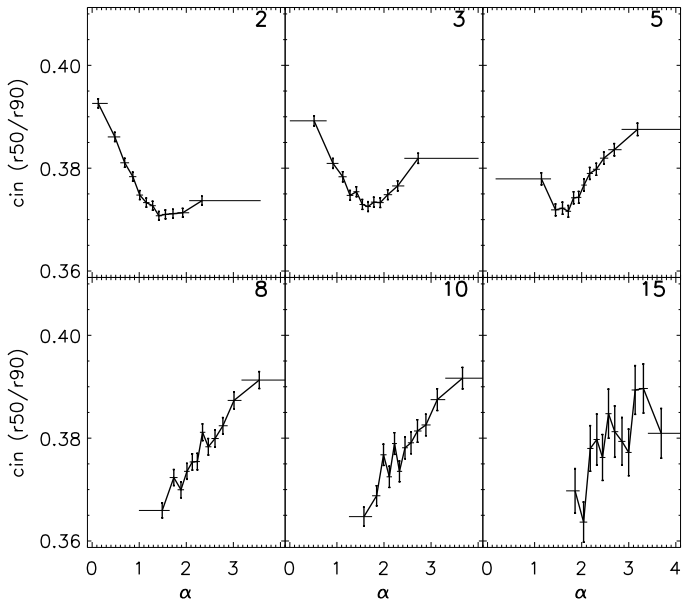

Figure 4.6: The mean concentration *cin* as a function of the mean geometry α at the given radii (upper left corner of each panel). In each subplot, the data is represented by 12 bins with an equal number of data points. The horizontal error bars of each bin indicate its width; the vertical errors of the y-axis the formal 1σ Poisson error.

are found in a point-like environment . Above the critical value $\alpha \sim 1$, no relation between *cin* and α is observed. In this regime, galaxies with a low *cin* (e.g. ellipses) have filamentary or uniform distributed neighbors within the radius, e.g. in a cluster.

At $r = 3\,h^{-1}$Mpc and $r = 5\,h^{-1}$Mpc , no global relation is observed, instead these scales seem to represent the transition from low small radius to the larger. The relation consists of the two slopes, the negative found at the smallest scale, and a positive, found at $r > 5\,h^{-1}$Mpc. It is interesting that these two radii are typical cluster scales. Beyond this scales, cluster are clearly identified because the sample volume includes all cluster member and not only a fraction of it. The relations at the intermediate radii $r \geq 8\,h^{-1}$Mpc is significantly correlated, where *cin* increases with increasing α. Clusters ($\alpha < 2$) are populated with galaxies of the lowest bulge-to-disk-ratio *cin*, void-like regions ($\alpha > 3$) with high *cin* galaxies (spirals). In between these two extrema of the α-*cin*-relation, there is a progressive incline. With this observed correlation, we can infer that galaxies with a geometry class between the two extrema (e.g. filaments) must have a intermediate bulge-to-disk-ratio between the extrema as well. This result indicates that galaxies in filaments are morphological different form cluster or void galaxies.

In order to distinguish between the morphology-density relation and morphology-geometry relation, the mean geometry α and the mean concentration index *cin* are plotted for a three density contrast δ_8 intervals in Figure 4.7. The average concentration index decreases for a constant radius from the low-density to high-density bin, which reflects the morphology dependence. In addition, the plot suggest that galaxies embedded in an environment with a similar density but different geometries can have different bulge-to-disk-ratios, which corresponds to different morphologies. In the lowest density contrast bin $\delta_8 < 0$, the concentration index decreases with increasing α. In the other two density intervals, the both galaxy properties are correlated for radii $r > 3\,h^{-1}$Mpc. Especially for galaxies the high density regions $\delta_8 > 2$, the geometry assigns different values which correlate with the concentration index. This may apply to the galaxies in filament-like regions that are located close the outskirts of cluster. Furthermore, this plot visualizes the scale-dependence of the geometry estimator for a constant density bin. For instance, the low-density bin (left column) mainly consist of galaxies in void-like regions, which have smallest α-values for $r = 2\,h^{-1}$Mpc and the largest values for $r = 15\,h^{-1}$Mpc.

4.3.2 Morphological Types

In addition to the presented plots, the concentration index can be used to group the galaxies sample into morphological types. In Figure 4.8, the fraction of spiral galaxies ($cin < 0\ 33$), mixed-type galaxies ($0\ 33 \leq cin < 0\ 375$), and elliptical galaxies ($cin \geq 0\ 375$) is plotted for local density contrast. Using this definition, the spiral galaxies are the most abundant morphological type. The grouping illustrates even more that the fraction of spiral galaxies declines with increasing density, while the fraction of elliptical galaxies increases. On all scales, the elliptical galaxies represent the largest fraction of galaxies in the highest density bin, while this morphological type represents the smallest fraction in the lowest density bin.

In Figure 4.9, the fractions of the three morphological types is plotted as a function of the local geometry. The fraction of spiral galaxies is highest in the void-like regions, decreases in the filament-like, wall-like and uniformly distributed environments, and reaches its minimum in the cluster regions. To separate the morphology-density relation from the morphology-geometry relation, similar relations were created for given density intervals in Figure 4.10. The density effect is again reflected in each line (constant radius) of this figure, where the fractions of spiral galaxies decrease and the fractions of elliptical galaxies increase from the low-density to the high-density interval. In addition to this environment dependence, relations between the morphology and the geometry can be found within a given density interval. For $r = 2\,h^{-1}$Mpc in the low-density interval and for $r \geq 3\,h^{-1}$Mpc, in the intermediate and high-density interval, the morphological types of galaxies depend on the geometry of their local environment of similar density.

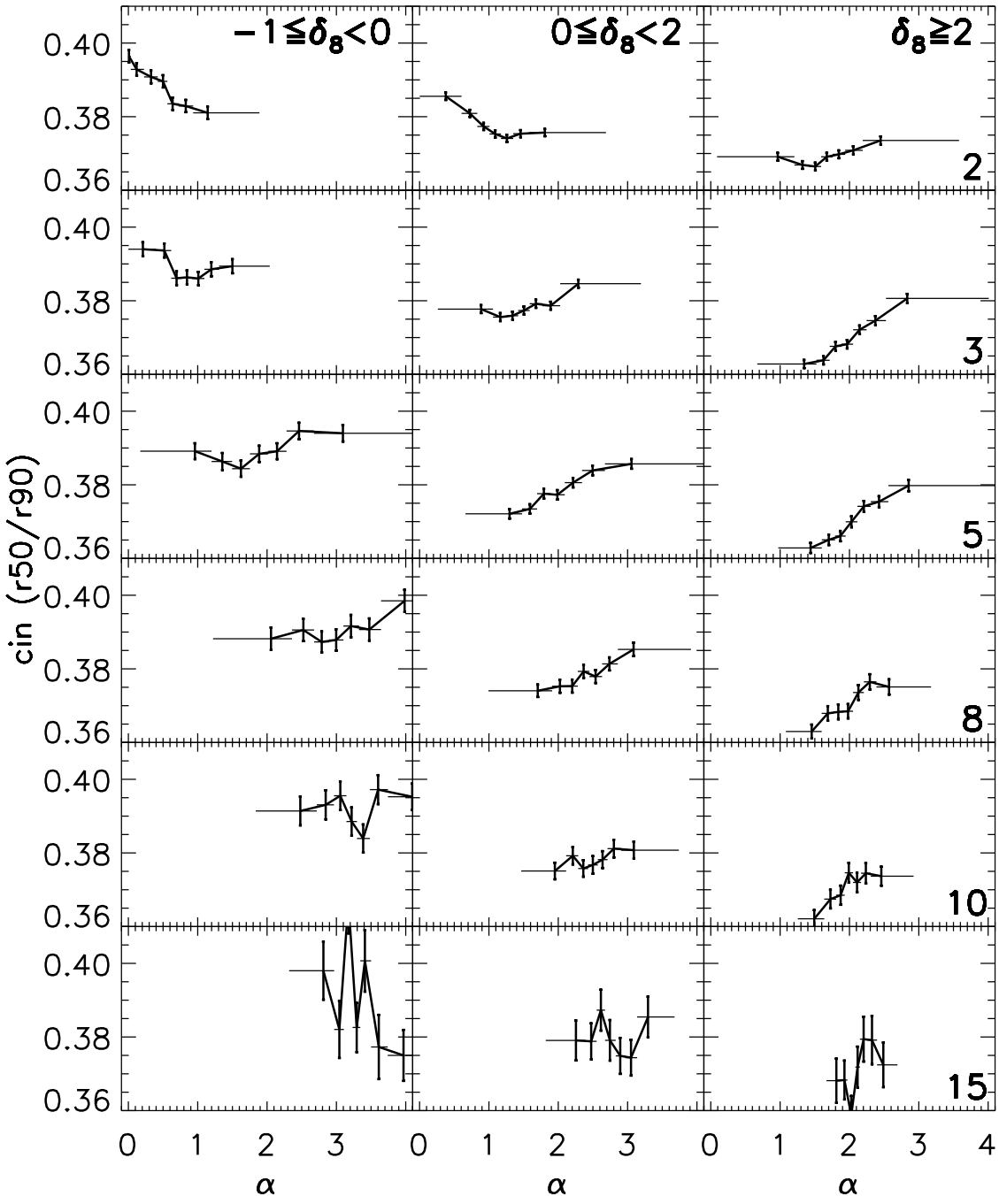

Figure 4.7: The mean concentration *cin* is plotted as a function of the mean geometry α for given density contrast δ_8 at $r = 8\,h^{-1}$Mpc (colums) and for different radii (rows). In each subplot, the data is represented by 7 bins with an equal number of data points. The horizontal error bars of each bin indicate its width; the vertical errors of the y-axis the formal 1σ Poisson error.

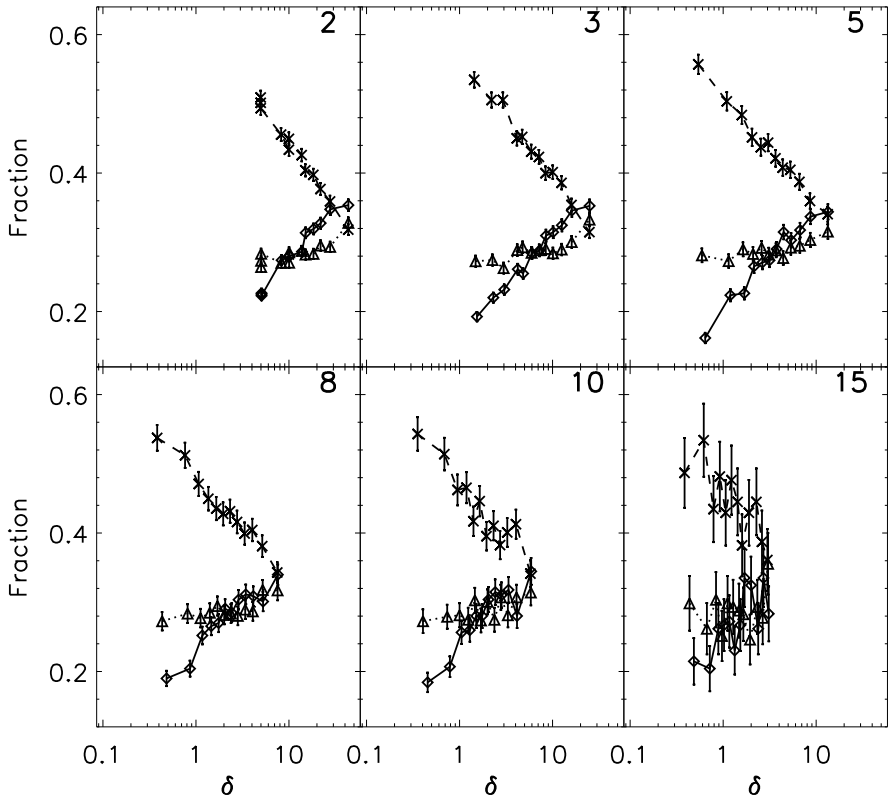

Figure 4.8: The fractions of morphological types are plotted as a function of the local density contrast. In each subplot, the data is represented by 12 density bins with an equal number of data points. The fraction of elliptical galaxies (diamonds, $cin < 0\ 33$), mixed-type galaxies (triangles, $0\ 33 \leq cin < 0\ 375$), and spiral galaxies (stars, $cin \geq 0\ 375$) is plotted for each density contrast bin. The vertical error bars indicate the formal 1σ Poisson error.

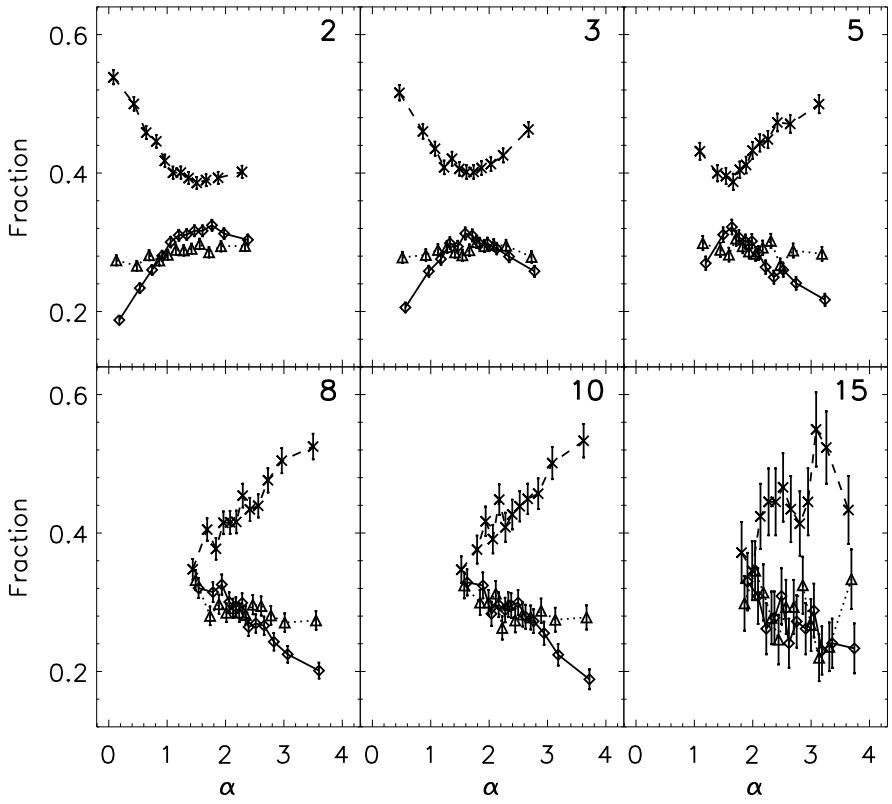

Figure 4.9: The fractions of morphological types are plotted as a function of the local geometry. In each subplot, the data is represented by 12 geometry bins with an equal number of data points. The fraction of elliptical galaxies (diamonds, $cin < 0\ 33$), mixed-type galaxies (triangles, $0\ 33 \leq cin < 0\ 375$), and spiral galaxies (stars, $cin \geq 0\ 375$) is plotted for each geometry bin. The vertical errors of the y-axis indicate the formal 1σ Poisson error.

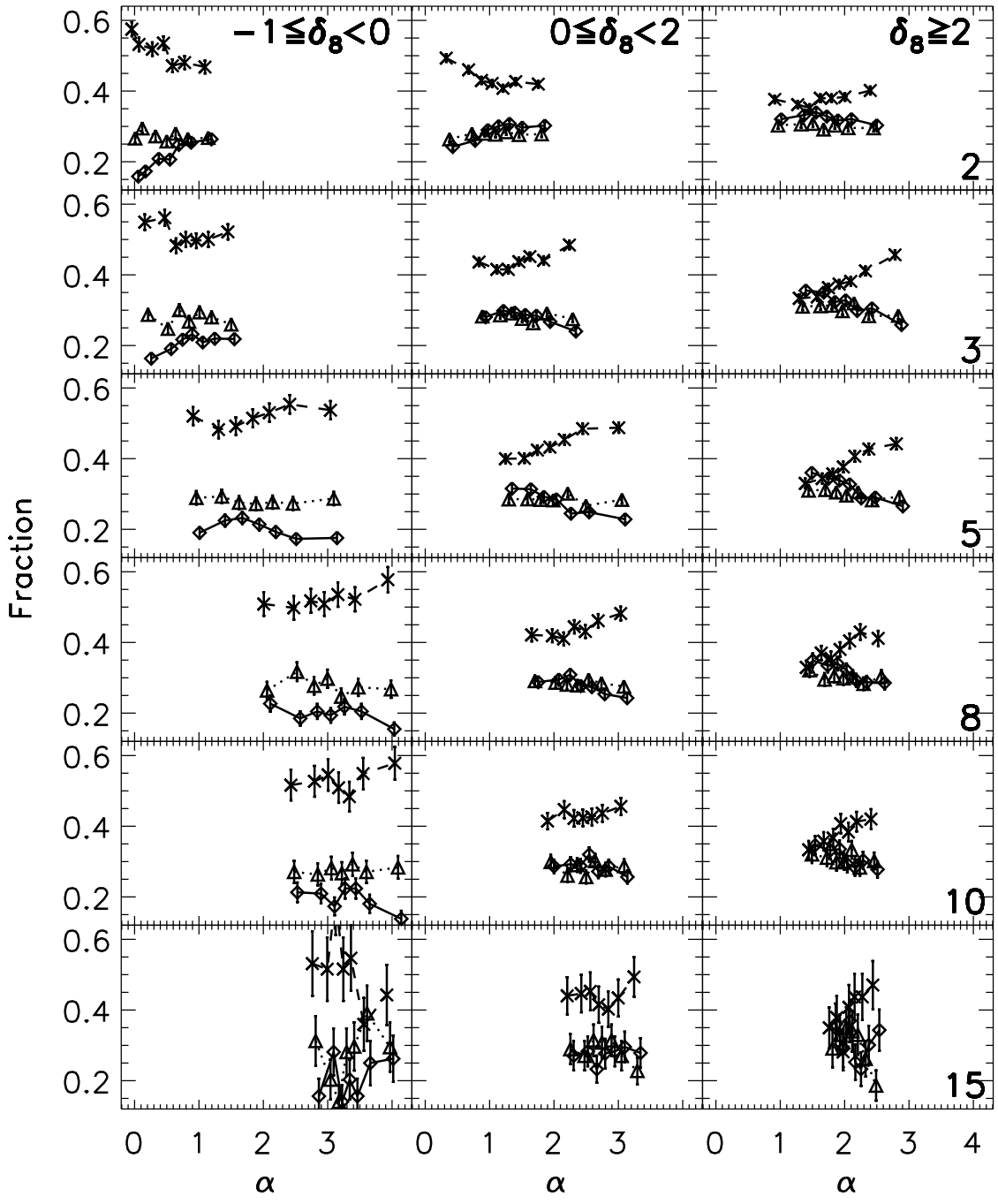

Figure 4.10: The fractions of morphological types are plotted as a function of the mean geometry α for three density contrast intervals δ_8 (at $r = 8\,h^{-1}$Mpc) (colums) and for different radii (rows). In each subplot, the data is represented by 7 bins with an equal number of data points. The vertical error bars correspond to the formal 1σ Poisson error.

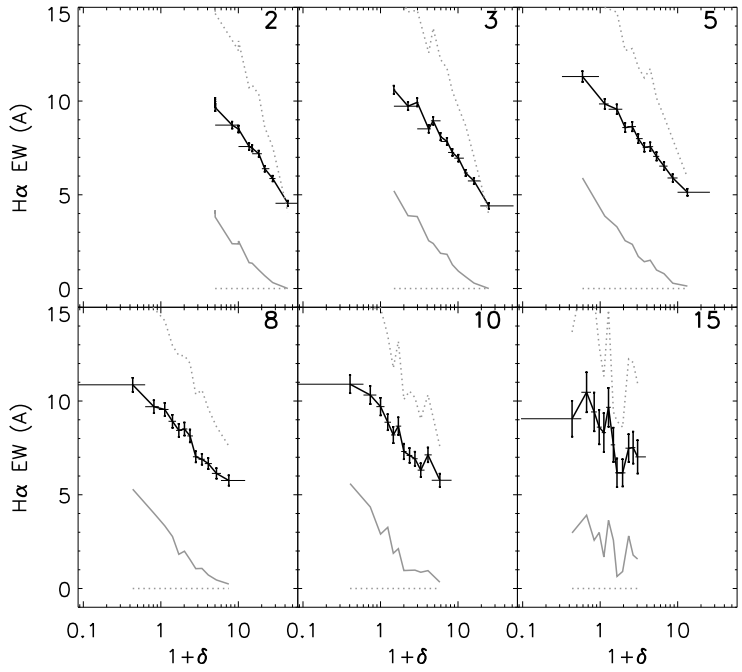

Figure 4.11: The mean Hα EW (black line) as a function of the galaxy density contrast δ at the given radii. In each subplot, the data is represented by 12 bins with an equal number of data points. The errors of the x-axis are the width of the bin; the errors of the y-axis the formal 1σ Poisson error. In addition, the median (gray line) and 25th and 75th percentile (gray, dotted lines) of the distribution are shown.

4.4 Star Formation

After the presentation of the photometric observations of galaxies, spectroscopic properties are plotted as a function of the local environment. For this purpose, the intensities of two emission lines that indicate the level of star formation were extracted from the SDSS data base.

4.4.1 Indicators of Star Formation

In Figure 4.11 and Figure 4.12, the mean intensities of the emission lines Hα and [OII], respectively, are both shown both as a function of local density contrast δ. Galaxies in low-density regions have lower intensities in both emission lines which indicates more star formation in those environments. The intensities decrease with increasing local density. Galaxies that are found in the tail of the Hα-distribution at EW$>$ 5 Å (Figure 2.8) have the broadest emission lines and strongly depend on the density, especially on the smaller scales $r \leq 5\,h^{-1}$Mpc. Slightly weaker but still significant relations were found for the [OII] emission lines.

In all density bins, the EW distributions are non-Gaussian which is shown by the inequality of the mean and median values. The EW-values are skewed to low values and have a long tail to large, positive values. The skewness is similar in each bin, but the high-density regions possess only a small fraction of galaxies ($\lesssim 25\%$) with high emissions EW$>$ 5 Å. It appears that low and normal star-forming galaxies are found in all density regimes, but that the strong star-forming galaxies are mainly found in low-density regions.

A difference between the Hα and the [OII] is found for the 25th percentiles lines which follow different slopes. The Hα 25th percentile is almost constant with the density contrast because the Hα EWs $\leq$ 5 Å are dominated and narrowed by stellar absorption. These effects make Balmer

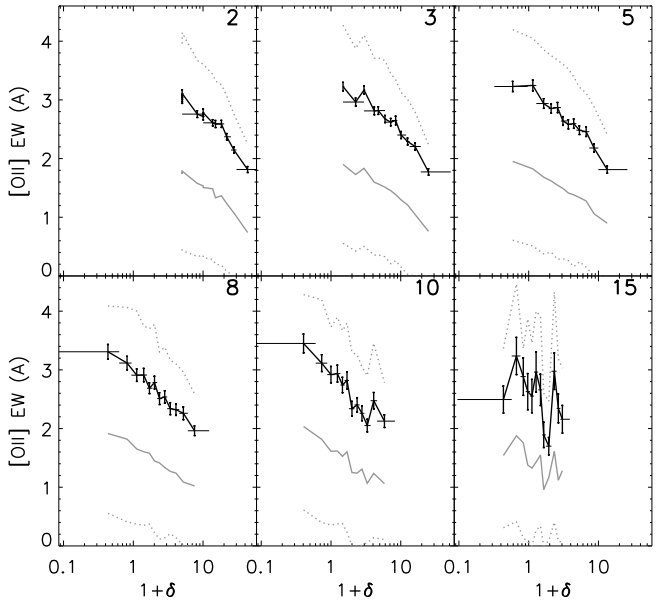

Figure 4.12: The mean [OII] EW as a functions of the galaxy density contrast δ at the given radii. In each subplot, the data is represented by 12 bins with an equal number of data points. The errors of the x-axis are the width of the bin; the errors of the y-axis the formal 1σ Poisson error. In addition, the median (gray line) and 25th and 75th percentile (gray, dotted lines) of the distribution are shown.

line only weakly sensitive to the stellar population of galaxies older than a few gigayears. The [OII] EW are unaffected by stellar absorption and reflect the common trend of declining emission line flux into the high-density regions.

After the relations with the density and the general comments, the emission line intensities of Hα and [OII] are plotted as a function of the local geometry in Figure 4.13 and Figure 4.14, respectively. Galaxies in cluster-like regions have the lowest intensities in the emission lines, while in the void-like regions galaxies with the highest intensities are found. In between, galaxies in filament-like, wall-like and uniformly distributed regions have intermediate intensities and, thus, may represent an transition between the low- and high-density regions, as well as an transition between region with high and low star formation.

On the smallest scale $r \leq 2\,h^{-1}$Mpc, the EW distributions are shifted to lower values for large α. Galaxies with the strongest Hα and [OII] emissions are mainly found in low-density, void-like regions which correspond to a $\alpha < 1$ The trend contiunes less pronounced in the $\alpha > 1$ regime, where galaxies with lower Hα and [OII] emission are found within a more populated environment as sampled with this small radius: filament-like regions, groups and clusters. The radii $r = 3\,h^{-1}$Mpc and $r = 5\,h^{-1}$Mpc, represent the transition scales, where the α-values of different geometries overlap. At the larger scales, $r \geq 8\,h^{-1}$Mpc, a low α-value corresponds to low Hα and [OII] emissions. Galaxies with a broad emission line populated structures with $\alpha > 2\,5$, i.e. uniformly distributed and void-like geometries. At $r = 15\,h^{-1}$Mpc, the highest geometry bin shows a drop of intensity in both emission lines. This drop of star formation for galaxies in large void-like regions may be explained by spiral galaxies that consumed all the gas in their local environment, so-called passive spiral galaxies.

In general, all geometry bins have different mean and median values of EW throughout the diagrams but both follow the identical overall trend. Both emission lines, Hα and [OII], have similar trends and resemble eachother regarding their dependence on the local geometry. Again, the main difference between these two emission lines is the 25th percentile, which is constant for Hα EW but follows the trend in the case of [OII]. As mentioned above, the measured Hα EWs

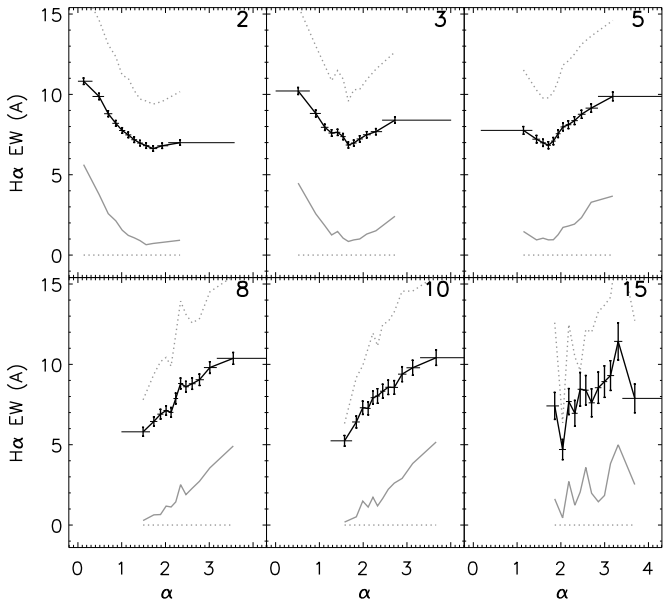

Figure 4.13: The black line represents the mean Hα EW as a functions of the local geometry α at the given radii in h^{-1}Mpc (upper left corner in each panel). Each bin contains 600 galaxies, the bin for the highest α contains 298 galaxies. The errors of the x-axis are the width of the bin; the errors of the y-axis the formal 1σ Poisson error. In addition, the median (gray line) and 25th and 75th percentile (gray, dotted lines) of the distribution are shown.

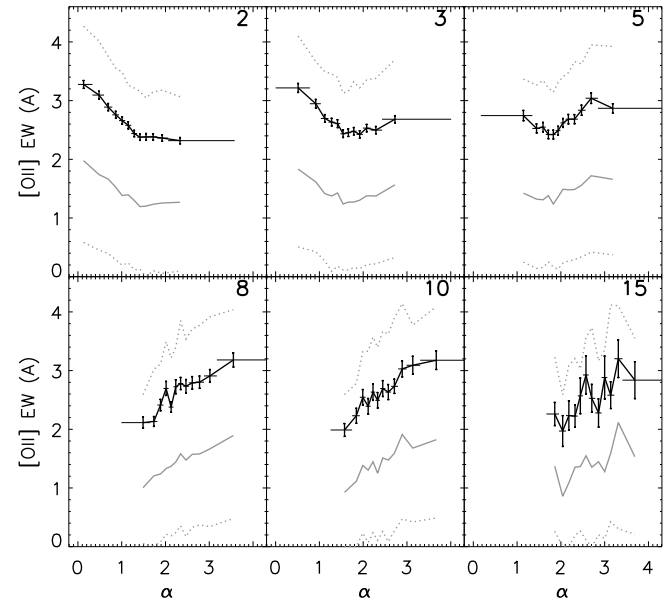

Figure 4.14: The black line represents the mean [OII] EW as a functions of the local geometry α at the given radii in h^{-1}Mpc (upper left corner in each panel). Each bin contains 600 galaxies, the bin for the highest α contains 298 galaxies. The errors of the x-axis are the width of the bin; the errors of the y-axis the formal 1σ Poisson error. In addition, the median (gray line) and 25th and 75th percentile (gray, dotted lines) of the distribution are shown.

are obscured by stellar absorption, whereas are the [OII] EW not.

Different environments can be identified through their typical α-value on each scale and each of these geometries exhibit different amount of emission line flux. To exclude the density effects in the relations, the emission line intensities are plotted as a function of the geometry for given denisity levels. In Figure 4.15 and Figure 4.16, the mean emission line EWs of Hα and [OII], respectively, are presented in this manner.

At $r = 2\,h^{-1}$Mpc, the low- and medium-density regions show a relation between the geometry and the intensity of both emission lines. The geometry estimator α distinguishs between different environments of similar low density and different star formation. Therefore, the geometry may indicate on-going galaxy evolution beyond the dense cluster regimes. On the scales between $r = 3\,h^{-1}$Mpc and $r = 10\,h^{-1}$Mpc, the intensities of the emission lines depend on the geometry in the high-density regimes, and slightly in the medium-density regimes. Galaxies in filament-like regions close to the outskirts of cluster may have similar densities like the cluster members but have a different local geometry. The largest scale $r = 15\,h^{-1}$Mpc, no significant trend was found. Again, each row illustrates the drop of the average emission line intensities from the low-density to the high-density interval, while the columns show the scale-dependence of the geometry estimator.

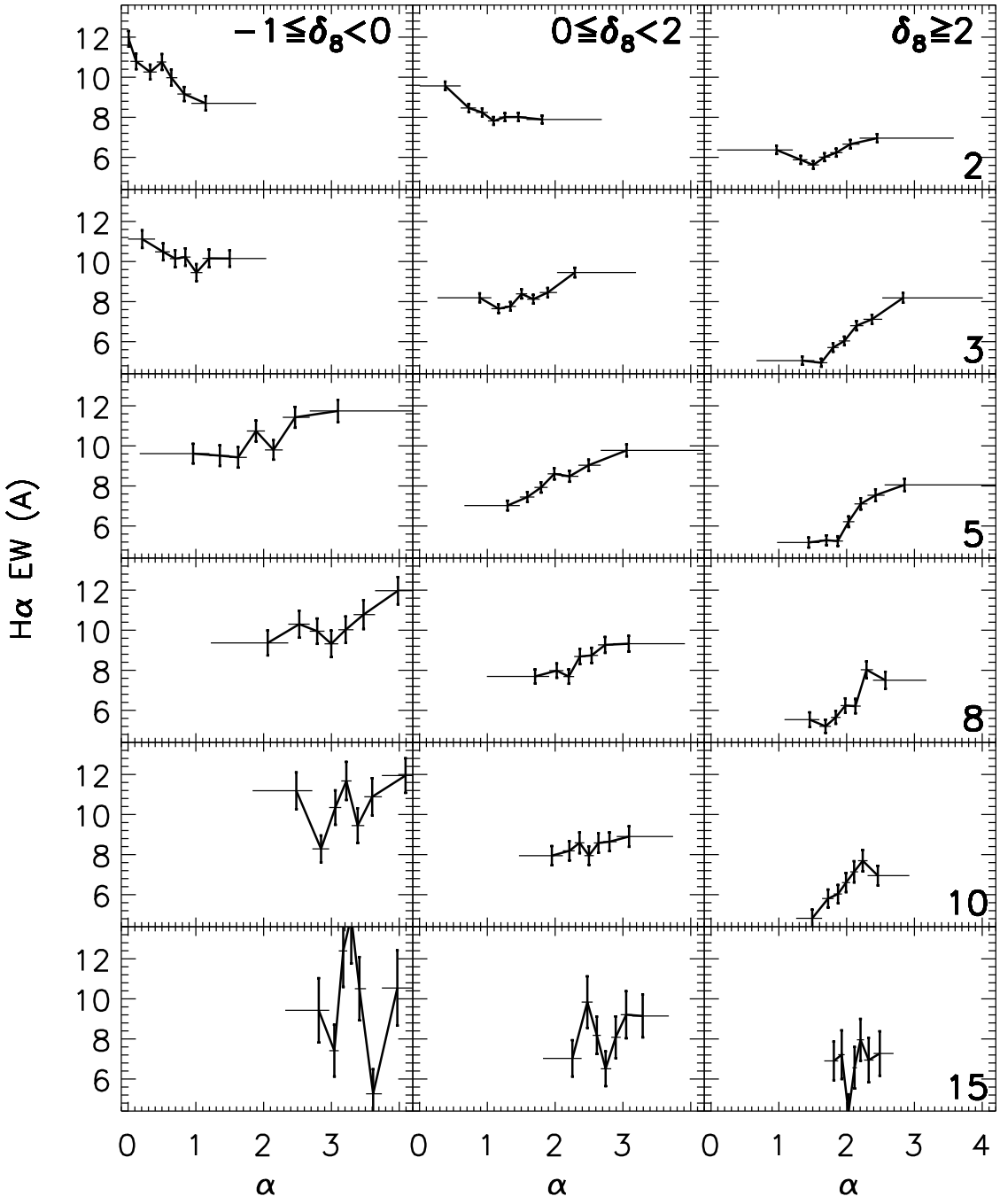

Figure 4.15: The mean Hα EW as a function of the local geometry α for given density contrast bin δ for three density contrast intervals δ_8 (at $r = 8\,h^{-1}$Mpc) (columns) and for different radii (rows). In each subplot, the data is represented by 7 bins with an equal number of data points. The horizontal error bars correspond to the width of the geometry bin; the vertical error bars to the formal 1σ Poisson error.

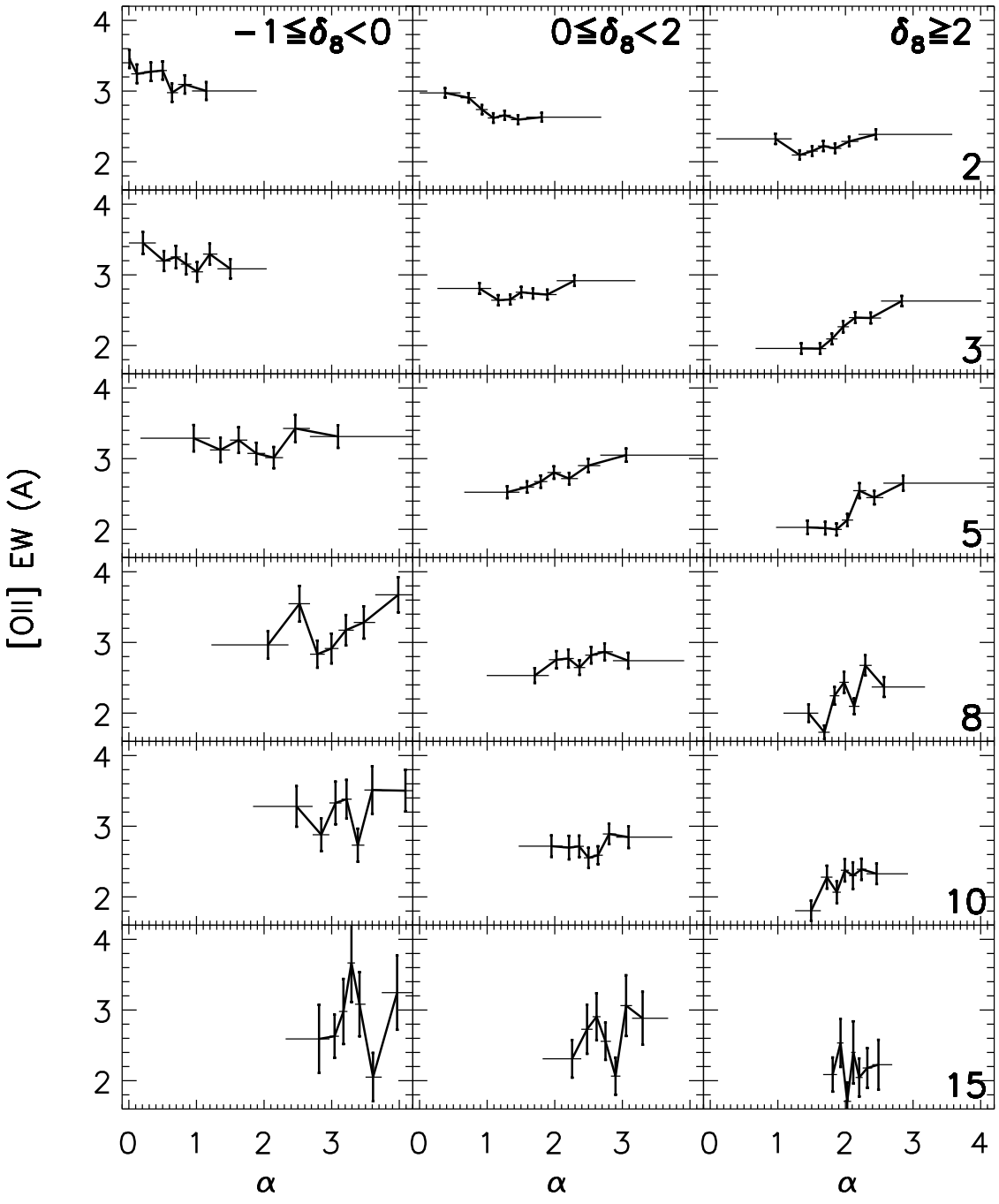

Figure 4.16: The mean [OII] EW as a function of the local geometry α for given density contrast bin δ for three density contrast intervals δ_8 (at $r = 8\,h^{-1}$Mpc) (colums) and for different radii (rows). In each subplot, the data is represented by 7 bins with an equal number of data points. The horizontal error bars correspond to the width of the geometry bin; the vertical error bars to the formal 1σ Poisson error.

4.4.2 Estimated Star Formation Rate

After the relations between the environment and the indicators of star formation, the emission line Hα is used to estimate the actual star formation rate (SFR). As mentioned above, the line Hα might be obscured by absorptions which can be corrected for each galaxy, as described in Section 2.7.

In Figure 4.17, the relation between SFR and the local density contrast is shown and confirms the results gained from the intensity Hα EW. On the small scales $r \leq 5\,h^{-1}$Mpc, galaxies in lowest density bin have the highest mean SFR of $\sim 0\,32\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$(median SFR $\sim 0\,14\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$). With increasing density, the SFR decreases to roughly half this maximal value for galaxies in the highest density bin (mean SFR of $\sim 0\,15\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$, median SFR $\sim 0\,05\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$). While for $r = 8\,h^{-1}$Mpc and $r = 10\,h^{-1}$Mpc, the relation is slightly reduced but still present and for $r = 15\,h^{-1}$Mpc , the relation between SFR and density is not significant anymore.

In Figure 4.18, the mean SFR is shown as a function of the local geometry. On the small scale $r = 2\,h^{-1}$Mpc and the scales $r = 8\,h^{-1}$Mpc and $r = 10\,h^{-1}$Mpc , the different geometries have different SFRs. For instance at $r = 10\,h^{-1}$Mpc , cluster galaxies have a mean SFR of $\sim 0\,18\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$(median SFR $\sim 0\,05\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$), and galaxies in filament-like, wall-like and unifromly distributed regions have a mean SFR of $\sim 0\,26\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$(median SFR $\sim 0\,08\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$), and galaxies in void-like regions have a mean SFR of $\sim 0\,32\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$(median SFR $\sim 0\,16\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$). Thus, galaxies in void-like regions have a two to three times higher SFR than galaxies in cluster region, and galaxies in filament-like, wall-like or uniformly distributed regions have roughly a 1.5 times higher SFR than cluster galaxies. The scales $r = 3\,h^{-1}$Mpc and $r = 5\,h^{-1}$Mpc represent the transition scales, and on the largest scale $r = 15\,h^{-1}$Mpc , the large error bars dominated the relation.

In order to separate the density effects, the estimated SFR is plotted as a function of the geometry for three density level in Figure 4.19. Relations between the geometry and SFR independent from the density were found on the smallest scale $r = 2\,h^{-1}$Mpc in the low-density and medium density interval, as well between $r = 3\,h^{-1}$Mpc and $r = 8\,h^{-1}$Mpc for almost all density intervals. In the low-density interval at $r = 8\,h^{-1}$Mpc, galaxies in void-like regions ($\alpha \sim 4$) have mean SFR of $\sim 0\,38\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$, and galaxies in filament-like and wall-like regions ($\alpha \sim 2$) have a mean SFR of $\sim 0\,30\,\mathrm{M}_{\odot}\mathrm{yr}^{-1}$. This may indicate that even in low-density regions $\delta < 0$, galaxies are found in environments with different geometries and those geometries are related with different SFRs.

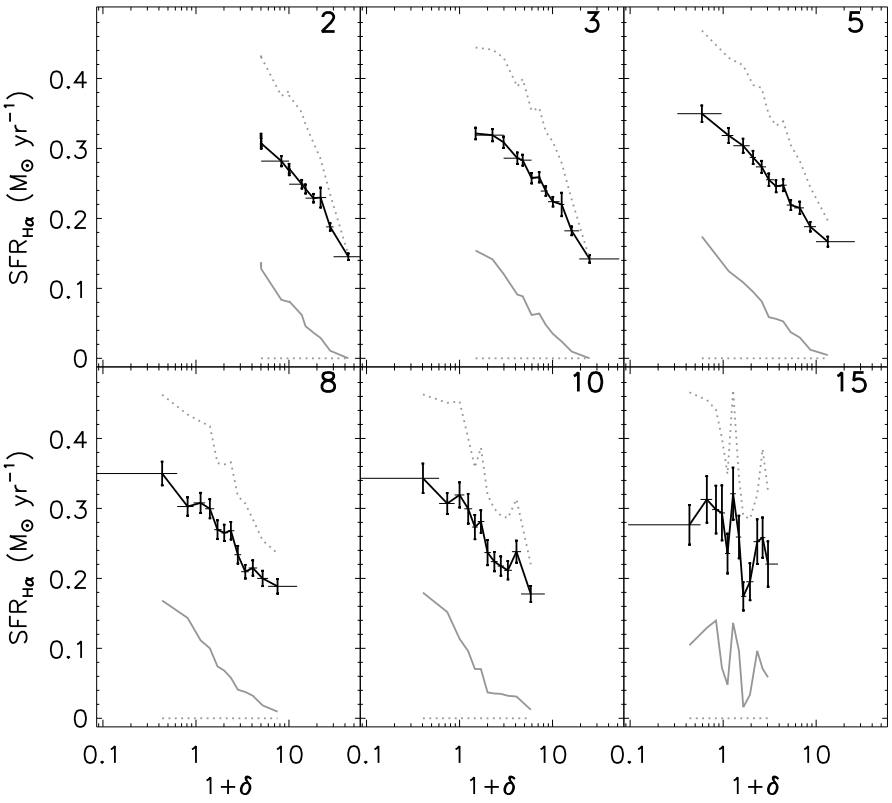

Figure 4.17: The mean star formation rate estimated from emission line Hα (SFR$_{\mathrm{H}\alpha}$, black line) is plotted as a function of the local density contrast δ at the given radii. In each subplot, the data is represented by 12 bins with an equal number of data points. The errors of the x-axis are the width of the bin; the errors of the y-axis the formal 1σ Poisson error. In addition, the median (gray line) and 25th and 75th percentile (gray, dotted lines) of the distribution are shown.

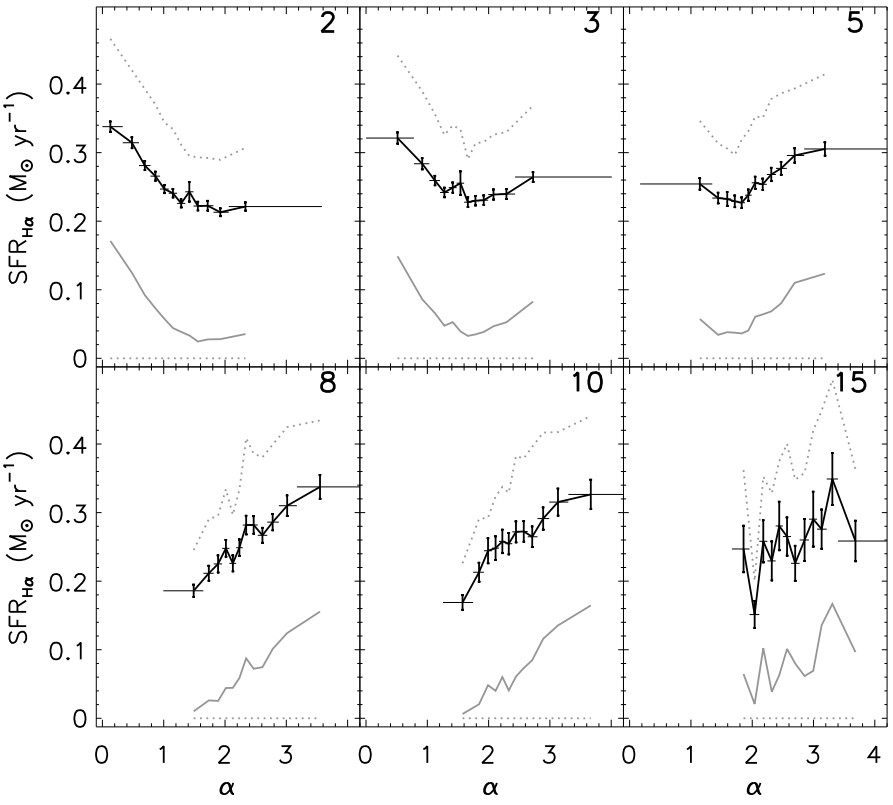

Figure 4.18: The mean star formation rate estimated from emission line Hα (SFR$_{\mathrm{H}\alpha}$, black line) is plotted as a function of the local geometry α at the given radii. In each subplot, the data is represented by 12 bins with an equal number of data points. The errors of the x-axis are the width of the bin; the errors of the y-axis the formal 1σ Poisson error. In addition, the median (gray line) and 25th and 75th percentile (gray, dotted lines) of the distribution are shown.

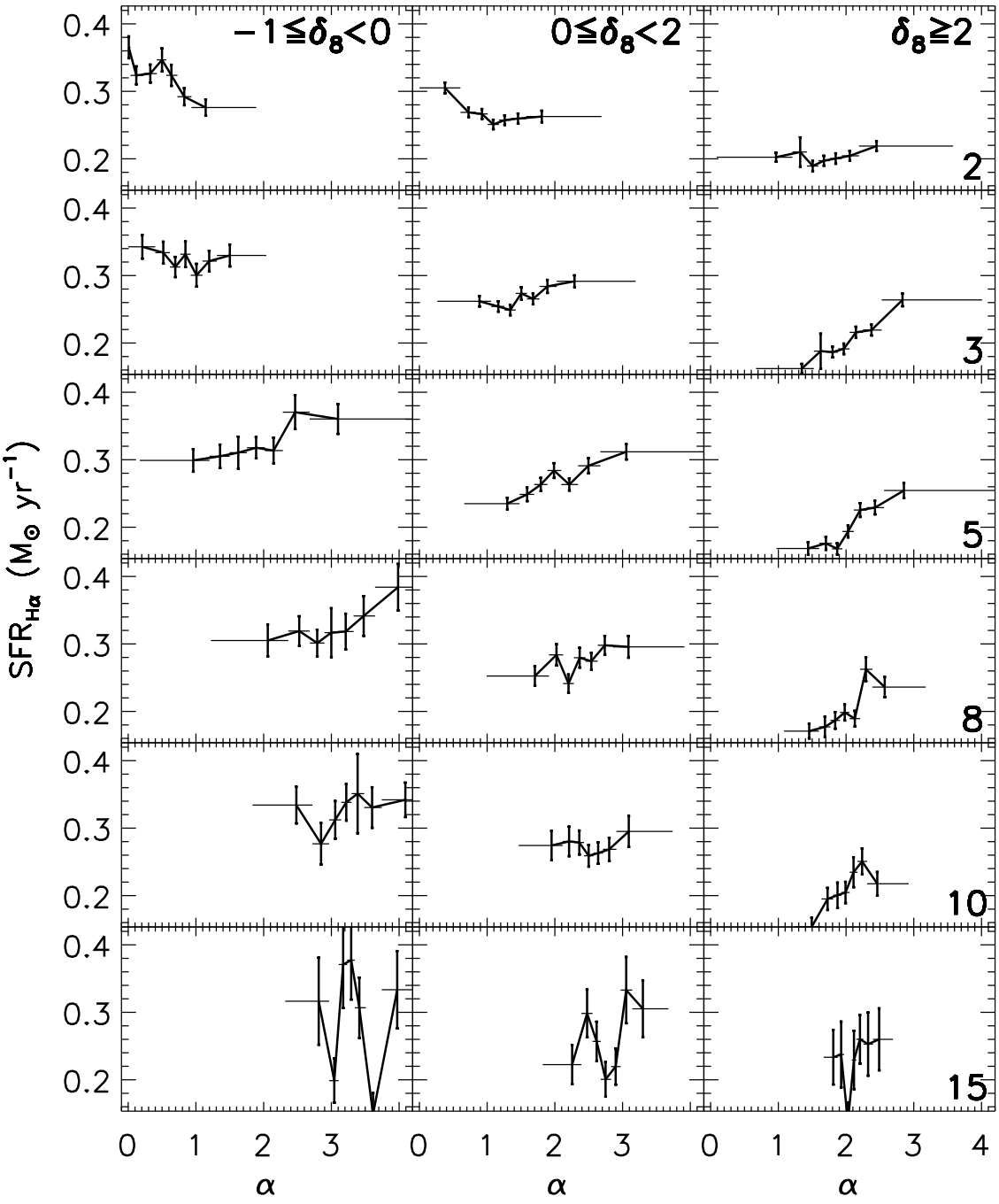

Figure 4.19: The mean star formation rate estimated from emission line Hα ($\mathrm{SFR}_{\mathrm{H}\alpha}$, black line) is plotted as a function of the local geometry α for three density contrast intervals δ_8 (at $r = 8\,h^{-1}$Mpc) (colums) and for different radii (rows). In each subplot, the data is represented by 7 bins with an equal number of data points. The horizontal error bars correspond to the width of the geometry bin; the vertical error bars to the formal 1σ Poisson error.

Chapter 5

Discussion

The Scaling Index Method was applied to estimate the local geometry of galaxy environments in a SDSS sample in order to understand the relation between local large-scale structure and galaxy properties. Compared with the local density, the geometry estimation provides complementary information by assigning each environment a specific value that corresponds to a certain shape, like clusters, filament-like, wall-like, uniformly distributed or void-like regions. Various relations were found between those local structure properties and other observable, physical properties of galaxies, namely e.g. color, morphology and star formation, which will have implications on the understanding of the cosmic large-scale structure and the evolution of the embedded galaxies. The following chapter discusses those implications, reviews possible sources of errors and compares the results with the literature.

5.1 SDSS Galaxy Sample

The SDSS was chosen for this work because it provides are large galaxy sample that can be used to apply new methods for the analysis of the large-scale galaxy distribution. The used flux-limited sample from the fourth data release (Adelman-McCarthy et al. 2006) included roughly $3\,8 \times 10^5$ from a total of $4\,8 \times 10^5$ galaxies with measured redshifts at a median redshift of $z_{0\,5} = 0\,10$. The sky coverage of the spectroscopic data set reached $\sim$4800 deg^2. This large sample size and volume allow stable estimations of statistical measures of the local galaxy environments. In addition, the SDSS catalog includes many examples of the typical constituents of the large-scale structure, like cluster regions that are connected with filament-like structures. One of the previous large galaxy surveys, the 2dF galaxy redshift survey (2dFGRS) (Colless et al. 2003; Cole et al. 2005), has a similar median redshift $z_{0\,5} = 0\,11$, but neither the large number of observed galaxies ($2\,2 \times 10^5$) nor the same sky coverage ($\sim$1800 deg^2). The 6dF galaxy redshift survey (6dFGRS) (Jones et al. 2009) has a larger sky coverage of $\sim$17000 deg^2, but a lower median redshift $z_{0\,5} = 0\,053$ and less galaxies ($1\,6 \times 10^5$).

For this work, it was important to use a large galaxy sample because the final analysis was performed with a strongly reduced data set. First, a volume-limited sample with $9\,4 \times 10^4$ galaxies was extracted from the flux-limited sample. This step was done to achieve a homogenous mean density within the sample and to exclude any redshift dependent selection effects. These selection effects may affect the calculation of the local geometry estimator α and the local density contrast δ and the resulting classification of the environment of galaxies may not be conclusive.

In a second step, galaxies close to the boundaries of the survey were excluded from the analysis to avoid any edge effects, which can lead to spurious classifications of the local environment. This approach of dealing with edge effects is called minus estimator or reduced sample estimator (Kerscher et al. 1998; Kerscher 1999). Without corrections, the geometries of galaxies at the boundary of the sample volume are biased towards wall-like environments because part of their

local environment is within the sample and the other part is outside. In addition, the local densities are underestimated as well. To calculate the local geometry and density without edge effects, the minus estimator was applied as a function of the radius. For larger radii, less galaxies remained in the galaxy sample and the variance of estimations increased. For the radii $r \leq 10\,h^{-1}$Mpc, the remaining galaxies were sufficient to retrieve statistical stable relations. Only for the largest radius $r = 15\,h^{-1}$Mpc, the small number of galaxies led sometimes to large error bars and, thus, to non-conclusive results (e.g. the morphology-geometry relation, Figure 4.6). An alternative approach of the minus estimator, which excludes many observed galaxies, might be to create an artificial galaxy distribution with specific statistical properties outside the survey boundaries. The edges and holes of the survey mask can be filled with randomly distributed data points with a density similar to the galaxy distribution (e.g. Rossmanith et al. 2009). For instance, the statistical moments of second order could be calculated from the existing galaxy distribution with the two-point correlation function and an artificial point distribution with the identical second order statistics can be created with e.g. a surrogate method (Raeth et al. 2002). However, the galaxy distribution cannot be completely described by the two-point correlation function, and thus cannot mimic the geometries that exist in the real galaxy distribution at low redshift values. This lack of information can introduce new source of errors and selection effects. Even with this approach, edge effects can still be present for those galaxies that are close to the boundary and include points from the artificial distribution outside the sample volume. To avoid those uncertainties, the minus estimator was used as a conservative measure to exclude edge effects, and, due to the large sample size of the SDSS galaxy data set, significant results were still obtained.

Another advantage of the SDSS data is the large amount of accessible photometric and spectroscopic measurements for each galaxy. For instance, whereas 2dFGRS only provides one band b_j, SDSS was designed to measure five luminosity bands, which can be used to define different color indices for the galaxies. With this data, it is possible to investigate various relations between the intrinsic physical properties of a galaxy and the structure information from the local environment. The modeling of the physical galaxy properties as a function of the new local structure properties like the scaling index will support the understanding of galaxy evolution and provides testable constraints for numerical models of galaxy formation. Connecting the local processes within a galaxy with the large-scale interaction with its neighbor galaxies is still one of the key questions in modern cosmology. The new structure statistic α was related with a full set of galaxy properties provided by the SDSS data and new insights were found that extend the well-known morphology-density relation.

5.2 Environment Classification

Instead of measuring the local density, the geometry estimator α provides complementary information and assigns each galaxy a geometry value that can be related to the shape of the local distribution of all neighbor galaxies within a certain distance. Typical geometries that are found in the large-scale galaxy distribution are cluster, filament-like, wall-like, uniformly distributed, and void-like regions. Those different constituents of the cosmic web were also found by other authors that applied different approaches, e.g. on the SDSS (Doroshkevich et al. 2004; Pandey & Bharadwaj 2006) and the 2dFGRS data (Pimbblet et al. 2004, see Section 5.3). In addition, the analysis of selected prototypes and toy models supported the existence of the different geometries and verified the ability of the introduced estimator α to distinguish between those classes of galaxy environments.

For galaxies in cluster and void-like regions, the density provides sufficient information because those geometries are defined by their local abundance of neighbor galaxies. In those galaxy environments, the geometry and density measures are related. However, for all the existing geometries in between these two extreme environment classes of low or high density, the geometry α is able to identify galaxies that are located in filament-like, wall-like, and uniformly distributed regions independent from the local density. Furthermore, it was shown that the geometry estimator dis-

tinguishes between environments of galaxies that have a similar density, but different physical properties, e.g. star formation rate.

This additional information that is offered by the geometry can be explained by its different approach to characterize the galaxy environment. The geometry α is a non-linear measure that estimates a local dimension of the surrounding galaxy distribution based on scaling relations within a given distance from the center galaxy. For instance, the environments of galaxies in filament-like regions have a local dimension $\alpha \approx 1$. This means that these environments can be roughly approximated by geometrical objects of a reduced dimensionality of 1, like lines or thin cylinders that include the neighbor galaxies. Environments of galaxies located in uniformly distributed regions have a local dimension of $\alpha \approx 3$, the dimensions of the configurations space, and cannot be represented by geometrical objects of lower dimensionality. In this view, the geometry α identifies those regions in the galaxy distribution that occupy volumes of a lower dimensionality than the configuration space. Individual galaxies in these lower dimensional volumes have had less spatial degrees of freedom during the large-scale structure formation, which not only can lead to a different local scaling behavior as measured by α today, but also to different local physical properties.

In contrast, the local density estimated with a top-hat filter just counts the neighbor galaxies within the given volume independent from the individual distances of each neighbor. Hence, same density values are assessed even if the relative distances of the neighbor galaxies inside the local volume were changed. For any center galaxy in a filament-like region, the local environment is not isotropically distributed and the local density will be underestimated because the average distance between the neighbor galaxies is actually lower. If the neighbor galaxies around the center galaxy were uniformly distributed, the density would be identical but the average distance between the galaxies would be larger. Unlike the geometry estimator α, the local density does not measure the local scaling behavior or the local dimension of the galaxy distribution. Thus, the density describes well the local mean distance between galaxies in cluster, void-like and uniformly distributed regions, while the geometry indicates the better description for galaxies in filament-like and wall-like regions, which cannot be defined by the local density alone.

The estimation of the local mean distance within the galaxy distribution reflects the local probability of galaxy interactions during the evolution of the large-scale structure. Because the absence or occurrence of different interactions can lead to specific physical properties of galaxies, the classification of the galaxy environments supports the understanding of galaxy evolution and can provide constraints for semi-analytic models. For instance, close galaxies share the same local gas reservoir and are more likely to merge, exchange gas, and to be affected by tidal forces. Those interactions alter the physical properties of the galaxies, e.g. by boosting the star formation rate that changes the color of a galaxy as well. In cluster and void-like regions, the local mean distance between galaxies is related to the local density and galaxy interactions. This explains e.g. the classical morphology-density relation (see Section 5.5), which is mainly found within and close to clusters, where the local density is strongly related with the influence of galaxy interactions like merging and ram-pressure stripping (Dressler 1980; Postman & Geller 1984; Goto et al. 2003a). In filament-like and wall-like regions, the local geometry is more related to both, the mean distance between galaxies and the probability of galaxy interactions, which leads e.g. to the observed color-geometry relation (Figure 4.4). These relations can result from interactions that still depend on the distance between two galaxies but do not require such extreme conditions, such as a high density of galaxies and hot gas that are found in galaxy clusters. Since the local density can not adequately estimate the probability of interactions for all constituents of the large-scale structure, the local geometry should be included to achieve a better description of the observed relations.

Whereas the geometry of galaxy environments indicates interactions during the evolution of galaxies, which depend on the mean distance, the question arises if the geometry can reveal information about the formation conditions of galaxies. In general, it is unclear if galaxy properties depend only on the environment that is observed today, or whether the mass of the dark matter halo during the formation of the galaxy determines the physical properties as well (Berlind et al. 2005). On small scales of about $\sim 1\,h^{-1}$Mpc, galaxy properties are more related to the local density and to the mass of the local dark matter halo. Larger radii sample a larger local environment

that includes more neighboring halos; this can reveal relations with the large-scale structure in which a galaxy is embedded. In addition, small halos host only one galaxy, whereas large massive halos include several galaxies that share an insufficient gas supply and whose star formation is truncated at an early stage. Larger halos result in cluster regions, which are well-described by the local density on small scales and their formation conditions. Thus, the local densities on small scales are more related with processes and conditions during the formation phase within large halos and with the galaxy properties that followed from those conditions. However, galaxies also emerged from smaller halos in the dark matter distribution, e.g. in filament-like regions, where the geometry provides a good environment description on scales larger than the halo size. Here, the geometry is less correlated with the local density and is more related with the evolution history that determines the properties of the galaxies. By measuring the observable galaxy distribution (baryons), one is measuring the underlying dark-matter distribution that includes filament-like structures as well. Because of the two different results from density and geometry, the observable relations depend on the scale and on the used measure to classify the local environment: if the local density is used on small scales of $\sim 1\,h^{-1}$Mpc, relations can be found that reflect mainly the galaxy formation within the large dark matter halos. If the local geometry is used on larger scales, the results reflect those galaxy interactions that have influenced the evolution of the galaxies since their formation.

5.3 Large-scale Structure Statistics

The picture that emerges from the results of the geometry estimator α shows a network of galaxies that are concentrated in clusters, filaments, and walls surrounding large empty voids. Although commonly used statistical measures like the 2-point correlation function (Section 3.1.1; Totsuji & Kihara 1969; Peebles 1973, 1980) provide comprehensive information about the scale, amplitude, and even the nature of the deviations from a uniform distribution, they at best yield only suggestive statistical measures for these structural patterns. Apart from the geometry estimator, other new approaches have been developed to analyze the structural features of the galaxy distribution in more detail.

The general topological analysis of the large-scale structure estimates the degree of connectivity of the galaxy distribution in the Universe. Once the redshift surveys have been smoothed with an appropriate filter, the topological features can be extracted from the isodensity surfaces. The results can be compared with theoretical distributions, e.g. Gaussian density fields, and non-linear properties can be detected. The topological genus statistic (Gott, Dickinson & Melott 1986) measures the relative abundance of overdense clusters to void regions and was estimated for the 2dFGRS (James et al. 2009) and a SDSS galaxy catalog (Gott et al. 2008), where on some scales slight deviations from the pure Gaussian field were found. An extension of this analysis, the Minkowski Functionals (Mecke, Buchert & Wagner 1994) have been calculated for galaxy surveys (e.g. Kerscher et al. 1998; Schmalzing & Diaferio 2000) like SDSS (Hikage et al. 2003) and galaxy cluster catalogs (Kerscher et al. 1997, 2001), where significant non-Gaussian features in the large-scale spatial distribution of galaxy clusters were found as well. Both the topological analysis and the geometry estimator α can detect non-linear structure components in the galaxy distribution. In contrast to the geometry estimator, these topological measures extract a global structure property for the complete galaxy sample. They can not be used to classify structure attributes of the local galaxy distribution that omits a direct comparison between galaxy properties and the galaxy environment. In addition, the geometry α does not only detect non-linear structures but allows to localize and label the corresponding galaxies, e.g. filament galaxies. The advantage of this approach is that the structure classification of individual galaxies can be related the their physical properties and that subsets of galaxies with a specific geometry can be created and studied, e.g. a set of filament galaxies.

A particular technique to quantify the filamentary character of the galaxy clustering is based on the minimal spanning tree (MST) that was introduced by Kruskal (1956). The MST is a

graph for a set of N points (here galaxies) that establishes a unique network of $N-1$ edges that link pairs of points. If there is a continuous path between all pairs of edges the graph is called connected; a connected graph with no circuits is called a tree. The MST is then the tree of minimal length that contains all points. In cosmology, MST has been used to determine the statistical significance of filament-like features. It was introduced by Barrow, Bhavsar & Sonoda (1985) as a statistic to find clusters and filaments in a sample of 489 galaxies selected from the CfA redshift survey. Using the same survey, Bhavsar & Ling (1988) demonstrated that the MST is a useful tool to show that filaments are real and not random alignments. In particular, they have shown that filament-like structures appear to radiate from the central regions of clusters. An effect that confirms the early view of Einasto, Joeveer & Saar (1980) that clusters were placed at the intersection of crossing filaments. These main insights about filaments could also be confirmed with the geometry estimator α. The histogram of the α distribution for the SDSS galaxy sample indicated an intermediate structure class of galaxies that are neither part of a cluster nor a void-like region. The investigated prototypes showed that the α values of these intermediate class corresponds to galaxies in filament-like structures.

Further applications of MST approach included the extraction of several quantitative distributions, for instance the number of edges per galaxy or the distribution of the edge length within the tree (van de Weygaert 1991; Krzewina & Saslaw 1996). Frequency histograms of the MST edge length were used by Doroshkevich et al. (1999) to analyze the galaxy distribution and to compare it with cosmological N-body simulations. In galaxy samples extracted from the Las Campanas Redshift Survey and SDSS, a similar analysis provided additional evidence for a network of structures that consists of a system of rich sheet-like structures, which in turn surround large underdense regions crossed by a variety of filamentary structures (Doroshkevich et al. 2001, 2004). In addition, the typical cell size of the filamentary network was estimated to be $\sim 10\,h^{-1}$Mpc.

Instead of analyzing the patterns of the galaxy locations, another type of cosmological statistics quantifies the shape and abundance of the prominent void-like regions within the galaxy distribution. This approach is motivated by an equivalence theorem by Cressie (1993, p. 625) proving that a point process can be completely described by its set of void probabilities. Since the complete shape of the void space cannot be quantified for observed or simulated data sets, the probability is calculated to find an empty sphere-shaped volume in the galaxy distribution, the so-called void probability function (White 1979; Ghigna et al. 1994; Rojas et al. 2004). It can be shown that the void probability function can be defined as an infinite series of n-point correlation functions of arbitrary order n (White 1979). Although this measure thus contains non-linear structure information like the geometry estimator α, it was shown that different realizations of dark matter distributions with an identical two-point correlation function were better distinguished with the geometry estimator α (Huber 2002). This can be explained by the spherical volumes of the void probability function that does not account for the true shapes and geometries of the observed void-like regions. In addition, the void probability estimates a global measure for the whole sample and does not classify the surrounding structure of individual galaxies. The geometry α combines the ability to detect galaxies in void-like regions, to measure non-linear structure components and to assign structure labels for each galaxy.

5.4 Luminosity and Colors

As shown in the results section, the different local geometry features of the environment of galaxies can be related to their physical properties like luminosity and color. The galaxy luminosity function (e.g. Blanton et al. 2001, 2003b) is a fundamental tool for the interpretation of observations and describes the number of galaxies per unit volume as a function of the luminosity. In SDSS and 2dFRS data sets, the calculated luminosity functions depend strongly on the local density of the environment: galaxies are more luminous in rich clusters (De Propris et al. 2003) and marginally more luminous in groups (Martínez et al. 2002) as compared to galaxies in the field (Madgwick et al. 2002). It was also shown that void galaxies are fainter than wall galaxies (Hoyle et al.

2005). These relations can be extended with the presented geometry estimator towards a more precise description of the local environment. Our results confirm the dependence of the luminosity on the environment and suggest that luminosity is related to the local geometry of the galaxy environment as well: cluster galaxies are more luminous than galaxies in wall-like regions and filaments and the faintest galaxies are found in void-like regions. This is in agreement with the observation that galaxies in filaments close to clusters are more luminous due to interactions that induce rapid star formation (Braglia et al. 2007).

Another interesting physical parameter of galaxies are their colors, which reflect the dominant star population in the galaxy and provide insights about the current star formation and star formation history of galaxies. Due to the observed bimodal shape of color histogram (Figure 2.5), a color threshold can be used the define two main groups in any galaxy sample: Red galaxies are dominated by old, metal-poor red giants, whereas blue galaxies consist of young, metal-rich stars. Although the observed galaxy colors correlate with the morphology (Humason 1936; Hubble 1936, Section 5.5), an advantage of the galaxy colors is the fact that they can be studied for more distant galaxies in the SDSS (Strateva et al. 2001; Baldry et al. 2004; Balogh et al. 2004b). While the morphological types can be assigned with some certainty only for nearby well-resolved galaxies, it is not possible for the fainter and more distant galaxies. If there is a relationship between the observed galaxy color and the morphological type, this relation may replace morphological segregation in studies of galaxy properties all the way to the limit at which one can perform a reliable star-galaxy separation.

Similar to the morphology, galaxy colors correlate with the galaxy density (Figure 4.3). It has been shown that blue and red galaxies populate environments of different density (Hogg et al. 2003; Balogh et al. 2004b). The red population of galaxies is a strong function of local galaxy density for fixed luminosities, increasing from 20% of the population in the lowest density environments to 70% at the highest, while the blue population show only a marginal dependence on the environment (Balogh et al. 2004b). Since the stellar population contains information about the star formation history and galaxy formation, the question arises what physical mechanisms led to this relation between the environment and the stellar population. A dominance of a red star population in a galaxy can be explained by an inability to form new stars owning to a lack of available interstellar gas, which has to be shared e.g. in high-density regions like galaxy clusters. As a consequence to the observations that both, the galaxy colors and the supply of gas, are dependent on the environment s galaxy density, one can also expect that the star formation rate of galaxies is a function of the environment as well (Section 5.6). This dependency results from the environment s influence on the stellar population during galaxy formation or due to interactions during the galaxy s evolution, or both. For instance, galaxies in high-density dark matter halos evolved fast, formed stars early and finished their gas supply earlier than field galaxies in the past which leads to the redder star population today. In contrast, blue galaxies in low-density regions formed later, evolved slower, formed stars at a lower rate, and have had larger gas supply (no neighbor galaxies), which is still sufficient to form young blue stars.

The observed color-geometry relation (Figure 4.4) is sensitive to evolutionary effects by augmenting the characterization of the environment by the local density with an estimate for the local geometry. The evolution of a galaxy can include different stages of e.g. star forming activity and galaxy-galaxy interactions that cause its physical properties. Those evolutionary stages can correspond and correlate with the different geometric types of environments a galaxy is passing through or is absorbed by. Parallel to the change and evolution of the stellar population of the galaxy, the environment of the galaxy can change due to its peculiar velocity caused by gravitational forces. However, it is suggested that those different environments can have similar galaxy densities but different geometry values in terms of how the neighbor galaxies are located relative to each other.

This scenario of a evolution of the galaxy environment s geometry may be of most importance for field galaxies, since they were formed outside of the larger dark-matter halos that were the origins for galaxy clusters. In this scenario, blue galaxies in void-like regions are attracted by filament-like structures or a cluster region, and eventually fall into them. During this transition,

the increasing number of close neighbor galaxies can induce a burst in the star formation rate and deplete the gas supply. As shown by the geometry estimator, galaxies in filament-like regions have properties that correspond to such a transition stage: brighter and redder compared to galaxies in void-like regions, but still less bright and bluer than cluster galaxies. This observation cannot be explained by a specific type of galaxy formation alone, because then there would be no environment dependency. Instead, certain processes during the galaxy evolution must be considered that influence the physical properties of galaxies in those environments. While the fate and the properties of a galaxy are determined by its formation in a large dark-matter halo together with many other galaxies, the properties of galaxies in void-like regions can change as well and can depend on their change of local environment.

5.5 Morphology

The color of galaxies is closely related to their morphological appearance that can be qualitatively described with the well-known Hubble sequence (Hubble 1926). Early-type galaxies (elliptical and lenticular morphologies) have a ceased star formation rate and consist of an old population of red giant stars. Late-type galaxies (spiral morphology) are characterized by a high star forming rate and consist of mainly young blue stars, which dominate the observed photometric properties. Those two main groups of galaxies dominate different environments according to the morphology-density relation (Dressler 1980; Postman & Geller 1984; Goto et al. 2003a), which was confirmed in this work (Figure 4.5), by estimating the morphology with the concentration index. Even early photometric surveys indicated that the fraction of early-type galaxies is higher in clusters of galaxies compared to low density environments (Hubble 1926; Oemler 1974; Dressler 1980). Based on these observations, the morphology-density relation appears to be an universal characteristic of galaxy populations (e.g. Postman & Geller 1984; Goto et al. 2003a; Smith et al. 2005). In quantitative terms, the morphological fractions are distributed over 3 magnitudes of projected galaxy densities, from the environments of cluster galaxies ($1000\,h^2\mathrm{Mpc}^{-2}$) to those of field galaxies ($\lesssim 10\,h^2\mathrm{Mpc}^{-2}$). Overall the morphology-density relation is slightly weaker than the color-density relation; while a change of morphology requires drastic interactions like ram pressure stripping (Quilis, Moore & Bower 2000), the stellar population is mainly controlled by the supply of interstellar gas.

A new finding in this work was the morphology-geometry relation (Figure 4.6) that indicates a correlation between the morphology of a galaxy and a quantitative measure for the geometry of its environment, the geometry estimator α. At a radius $r = 10\,h^{-1}\mathrm{Mpc}$, the morphological mix of galaxies is dominated by high fractions of elliptical galaxies for small geometry values (cluster-like regions), while the largest geometry values (void-like regions) are dominated by spiral galaxies. In between those two regimes, a transition was found in terms of both, morphology and geometry. The fraction of spirals increases with increasing geometry values from cluster-like regions to filament-like and wall-like regions, and reaches the highest values in the void-like regions.

Both, the observed morphology-density and the morphology-geometry confirm the local environment s influence on the morphology of galaxies, which is based on the effects of interactions between neighbor galaxies and the supply of interstellar gas. Late-type galaxies (high concentration index) are mainly found in under-dense regions, where the evolution of the disk is barely disturbed by any kind of interaction with neighbor galaxies or the cluster gas. High density regions consist of galaxies with low concentration indices, like elliptical galaxies. In these environments, galaxies were inhibited to form a disk due to the cluster collapse in the early stages of their evolution, the hot inter-cluster medium that prevented star formation, and the high probability of interactions with close neighbor galaxies. The importance of this interaction probability is also shown by the relation between the morphology and the distance to the cluster center, the morphology-radius relation, which was found to be more fundamental than the morphology-density relation (Whitmore & Gilmore 1991; Whitmore et al. 1993). The cluster environment slows down the formation of the disk of close galaxies and increases the amount of hot gas in the cluster leading to a correlation

between the fraction elliptical galaxies and its X-ray properties (Whitmore et al. 1993). Also, S0 galaxies are considered as remnants of spiral galaxies whose gas was stripped away or evaporated (Dressler 1980).

Compared to the local density, the geometry estimator α can provide new insights in the relation between morphology and environment by identifying specific types of galaxies in e.g. filament-like environments. In some cases, the connection between galaxy color and morphology cannot be easily explained with the model mentioned above. Some spiral galaxies have been observed to have a low concentration index, an indication of an elliptical morphology, because the star formation in the disk ceased due to gas shortage. Such spiral galaxies with none of the typical emission lines, so-called anemic galaxies (van den Bergh 1976), passive spirals (Couch et al. 1998) or k-type spirals (Dressler et al. 1999), are found in the in-falling regions of clusters. Goto et al. (2003a) suggested that passive spirals represent an intermediate stage in the galaxy evolution between the spiral and elliptical morphology. Numerical simulations indicate that passive spirals can provide valuable information about less dramatic but long-term environmental effects on galaxy evolution (Bekki, Couch & Shioya 2002). The existence of passive spirals can explain an intermediate morphology class of galaxies in filament-like or wall-like regions as found by the morphology-geometry relation. These results indicate that galaxies in filament-like and wall-like regions are morphological different from galaxies in cluster or void-like regions and can be regarded as a transition class between those two extreme density environments. Furthermore, not only the the number of close neighbors (local density) seem to be important to the evolution of galaxies, but also how these neighbors are distributed (local geometry).

The morphology-geometry further augments the classic morphology-density relation because the observations and implications of environmental effects on galaxy properties are extended to lower density regions. Studies investigating the morphology-density relation mainly include cluster galaxies and galaxies close to clusters (e.g. Dressler 1980; Postman & Geller 1984; Goto et al. 2003a), whereas the SDSS sample used in this work contains a representative set of large-scale structure elements of the galaxy distribution including geometrically different low-density environments. These environments have not been considered yet to have any effect on galaxy evolution or formation, because the previously used density measures show no correlation with different galaxy properties beyond the outskirts of clusters. In this study, a correlation was found for even low-density regions because the geometry of galaxy environments provides additional information that reflects a local probability of interactions. The possible process which accounts for this relation must be different form the well-known interactions with hot cluster plasma, like ram-pressure stripping or gas evaporation (Gunn & Gott 1972; Dressler 1980; Abadi, Moore & Bower 1999; Quilis, Moore & Bower 2000) due to the lack of high densities. Instead, a less dramatic but long-term effect on galaxy evolution must be considered instead, for instance tidal interactions (Nikolic et al. 2004). The existence of passive spirals points out that those mild but persistent processes are indeed possible and responsible for the properties of some galaxies in low-density regions.

Beyond these regions, the influence of environments of distinct geometries can be imprinted in the physical properties of galaxies. During their evolution, galaxies can be members of different environments changing and adapting their properties according to the properties of the region. A possible scenario for an evolution history can be the formation in a low-density region, to drift into a filament-like environment, and then being pulled into a cluster region. The intermediate density of filament-like regions is between the extreme densities of void-like and cluster-like regions, and thus correlates with the intermediate probability of interactions with neighboring galaxies: galaxies in filament-like regions are likely to have more interactions than in void-like regions and less interactions than in clusters. The morphology-geometry relation suggests that filament galaxies belong to a intermediate transition environment with the corresponding intermediate mix of morphological types like spiral and elliptical galaxies.

A similar effect of environment transitions is known for the morphology-density relation: galaxies are pre-processed in galaxy groups before they are pulled into the cluster regime. Caused by this intermediate state before entering the cluster, the morphology-density relation is even observed in the outskirts of clusters beyond its viral radius. This argument further supports the intermediate

morphology of filament galaxies, since the ends of filament structures are connected to clusters and flow into them due to the gravitational attraction.

5.6 Star Formation

The observation of the galaxy s morphology requires the existence of luminous stars that indicate the baryonic matter distribution within the galaxy. Since the morphology depends on the density of the environment, the question arises whether the star formation rate (SFR) of a galaxy depends on environment properties as well. There is no definite separation between this dependency from other relations, since the morphological type and the environment are strongly correlated (Dressler 1980; Postman & Geller 1984; Goto et al. 2003a), as well as the morphology and the star formation (e.g. Kennicutt 1998a). Since the SFR is crucial for the understanding of the galaxy evolution, it is interesting to quantify the influence of the environment specified by local density and local geometry estimations.

With the local galaxy density, a strong relationship between the SFR and environment was found in this work. On scales $r \leq 10\,h^{-1}$Mpc, the SFR indicators, the distributions of Hα and [OII] EWs, changed as a function of the density. Especially the 75th percentile of those distributions, i.e. strongly star-forming galaxies, show a strong decrease with increasing density. This observation corresponds with results from Gómez et al. (2003), who found that the distributions of Hα and [OII] EWs change as a function of the local projected galaxy density using the SDSS Early Data Release. They characterized the effect in three ways: a shift in the overall distribution to lower SFR values with increasing density, strongly star-forming galaxies (Hα EW > 5 Å) that are noticeably decreased in high-density regions, and a characteristic scale in the correlation between SFR and density at a galaxy density of $1h_{75}^{-2}\mathrm{Mpc}^{-2}$. Consistent results were observed by Hashimoto et al. (1998), who investigated the influence of environment densities on the SFR in the Las Campanas Redshift Survey, and by Lewis et al. (2002), who studied the same relations in the 2dF galaxy sample. Both groups were able to identify a critical density and radius, where the SFR of cluster galaxies changes from a specific SFR for field galaxies. At the turnaround radius $R_t \quad 5R_v$ with the cluster s viral radius R_v, the SFR becomes statistically identical to the field. Those results are in agreement with numerical simulations suggesting that the main physical properties of the SFR is controlled by the amount of cool gas and the time since the last interactions with a larger halo. For instance, Diaferio et al. (2001) predicted that the mean SFR should be lower than the field out to $2R_v$. In addition, the results of Gómez et al. (2003) are qualitatively consistent with hierarchical models showing that structure formation can affect the SFR of galaxies beyond the virial radius. Future studies can show if simple models are adequate or additional physical models are needed to explain the effect on the SFR in galaxies that are located in more distant cluster infall regions between $2R_v$ to $4R_v$.

With the local geometry, a novel relationship between the SFR and the environment was found even outside high-density regions (see Figures 4.13, 4.14, 4.18). For scales $r \leq 10\,h^{-1}$Mpc, the used SFR indicators are functions of the geometry estimated by the scaling index α. The lowest SFR was observed for galaxies in cluster-like environments, whereas galaxies in filament-like and wall-like environments showed higher SFRs. Galaxies in void-like environments had on average the highest SFR. Similar to the results obtained with the local density, the strongly star forming galaxies showed the highest sensitivity in respect to the geometry of galaxy environments. In order to exclude the effect of the density, these relations were also plotted for galaxies with similar local density. It was shown that even in this case, relations between the local geometry and the SFR indicators were obtained (see Figures 4.15, 4.16, 4.19). These results indicate an existing correlation between SFR and the local geometry that is partially independent from the local density. Since various different local geometries constitute the large-scale structure of the galaxy distribution, these relations can be used to formulate constrains for numerical simulations and models of galaxy evolution.

The continuous transition between the two extreme regimes, a high SFR in void-like environ-

ments on the one side and a low SFR in cluster-like environments on the other, can be explained by a simplified, yet plausible scenario for galaxy evolution and indicates a correlation between star formation history and geometry history of a galaxy. Typical spiral galaxies arise in low-density regions, aggregate gas from the interstellar medium and form blue stars. They are attracted by close elements of the large-scale galaxy distribution such as filament-like environments, which flow towards galaxy clusters. When such early-type galaxies enter a cluster region, different mechanisms like ram pressure stripping or harassment deplete the galaxies gas and transform them into a late-type galaxy. During this evolution path through the different environments, the SFR can be altered resulting in a specific signature in the star formation history. If only the well-known effects of the cluster environment alter galaxies properties, only cluster galaxies would have a different SFR compared to other galaxies in the field. Then, the relation between environment and SFR would be similar to a step function, where the SFR changes at a critical geometry from a typical value for field galaxies to a typical value for cluster galaxies. Instead, the smooth and continuous transitions in relations between the geometry and the measured SFRs (Figure 4.13 and Figure 4.14) indicate that galaxies in filament-regions have different SFRs compared to both, cluster-like and void-like regions.

This observation indicates that the SFR can be changed in filament-like regions by interactions between galaxies and physical processes. One possible source of mutual influence can emerge from a burst of star formation induced by tidal interactions between close galaxy pairs (Nikolic et al. 2004), which are more likely in filament-like regions. It was shown in numerical simulations that weak bulges in spiral galaxies facilitate the creation of bars during tidal interactions (Mihos & Hernquist 1996). The bar allows gas to flow towards the galaxy s nucleus. The following induced star formation continues even after the galaxies are widely separated. In contrast, a strong bulge stabilizes the galaxy and inhibits bar formation, which leads to less star formation. Another reason for the observed relations can be the smaller gas supply in filament galaxies. Usually, more interstellar gas is available around galaxies in void-like regions compared to cluster regions. Since galaxies in filament-like regions share the gas reservoir with close neighbors, their gas supply is more limited compared to galaxies in void-like regions, but more abundant compared to cluster galaxies. The limited gas supply can lead to reduced star formation. It is also possible that the star formation ceases if the gas supply drops below a certain threshold (e.g. Madau, Pozzetti & Dickinson 1998; Bruzual & Charlot 2003).

Chapter 6

Summary

The Scaling Index Method (SIM) was applied on a volume-limited SDSS galaxy sample in order to investigate the correlation between the local geometry of galaxy environments and the physical galaxy properties. With this new application of the SIM approach, it is possible to assign each galaxy a characteristic number that corresponds to e.g. a uniform, filament-like, wall-like, cluster-like or void-like sub-region within the galaxy distribution. This local geometry characteristic extends and complements previously used local density estimations by not only counting the occurrence of local galaxy neighbors, but also considering the shape of the local galaxy distribution in terms of scaling relations. By assigning each galaxy a SIM value, different classes of galaxy environments were identified and classified in order to compare their physical properties. Such a representation provided relevant insights into the connection between the large-scale structure of the galaxy distribution and the understanding of evolution histories of individual galaxies that induced several new relations between these properties.

Specifically, it was shown that the local scaling properties estimated by this new SIM approach correlate with the physical galaxy properties like luminosity, color, morphology, and indicators of star formation. Apart from the well-known dependence of these galaxy properties on the local galaxy density, e.g. morphology-density relation, the results indicate a new dependence on the shape of the local galaxy environment, which was expressed by the morphology-geometry relation. While the initial conditions during the galaxy formation, e.g. the halo mass, have a strong influence on galaxies especially in galaxy clusters, the results show that galaxies outside high-density regions can still be affected by the environments during their evolution, e.g. in filament-like structures. Different types of environments, as categorized by SIM, influence the residing galaxies differently, e.g. due to the local galaxy-interaction probability or by providing different levels of gas supply. This dependence of the galaxy properties on the local geometry remained even for environments with constant densities, which indicates that local geometry extends and complements the previously used density-based structure descriptions of galaxy environment classes that are relevant for the galaxy evolution.

It was shown that the estimation of the local geometry using the SIM approach offers several advantages for the analysis of observed galaxy distribution. Compared to other geometrical or topological measures, it calculates a local structure index for every galaxy, which is useful for the analysis of galaxy properties. Paired with the clear interpretation of SIM, the structure information can be easily visually confirmed, if necessary, and constrained subsamples of certain geometry classes can be created. For this purpose, the interpretation of certain SIM values was demonstrated with simplified structure models that mimic certain structure elements in the galaxy distribution, e.g. filament-like and wall-like environments. The stability of the geometry estimation was analyzed and confirmed by considering the effects of possible sources of errors, including intrinsic redshift distortions, and by using numerical simulations of the SDSS galaxy distribution.

Possible extensions of the SIM approach include the use of a non-isotropic distance measure and the analysis of more galaxy properties like metallicity or different types of surveys, e.g. quasars. A

further application of the SIM approach can be the analysis of numerical simulations of the galaxy distributions, where the correlations between galaxy properties and geometry can be compared with relations found in galaxy surveys. In addition, the analysis of the local geometry as a function of time for individual galaxies can address questions about what environments are occupied during the galaxy evolution and how different evolution histories lead to different galaxy properties observed today.

Appendix A

Additional Resources

A.1 Derivation of the Geometry Estimator from the Theory of Fractals

The two-point correlation function of galaxies has a nearly constant slope over a wide range of distances (see Section 3.1.2). This self-similar behavior for galaxy clustering suggests that the concept of fractal geometry may apply (Mandelbrot 1982; Grassberger & Procaccia 1983; Peebles 1993). Various fractal analyses have been applied, such as wavelet transforms (e.g. Martinez et al. 1993) and percolation analysis (e.g. Klypin & Shandarin 1993). In this Appendix, we use this self-similar behavior to provide an alternative derivation of a geometry estimator, the scaling index method.

Consider a given point set $\mathcal{S}$ which is a fractal object and can be defined on a natural probability measure $\mathrm{d}\mu$. Then, there is a decomposition into the subset $\mathcal{S}_\alpha$:

$$\mathcal{S} = \bigcup_{\alpha} \mathcal{S}_\alpha \tag{A.1}$$

The variable α is a continuos variable which describes the local scaling properties of the fractal object $\mathcal{S}$ and can be considered a local scaling exponent:

$$\alpha_i = \lim_{\epsilon \;\; 0} \frac{\log p_i(\epsilon)}{\log \epsilon} \tag{A.2}$$

where $i = 1 \qquad N(\epsilon)$ and $N(\epsilon)$ is the number of necessary cubes with the edge length ϵ to cover point set $\mathcal{S}$. We define the probability $p_i(\epsilon)$ as

$$p_i(\epsilon) = \int_{\mathcal{K}_i} \mathrm{d}\mu(x) \tag{A.3}$$

estimates the number of points in the cube $\mathcal{K}_i$ at the location x. The scaling exponent α_i is the slope of $\log p_i(\epsilon)$ as a function of $\log \epsilon$ as describes how fast the points in the cubes decrease with the decreasing ϵ. In case of $\alpha_i = \text{const}$ for all i, $\mathcal{S}$ is a mono-fractal. If two slopes are found, $\mathcal{S}$ is a bifractal, and if even more scaling exponents are found, we call $\mathcal{S}$ a multifractal.

In most applications, the point set under consideration is finite and the the limit $\epsilon \qquad 0$ in Eq. (A.2) is not defined. For these cases, Grassberger et al. (1988) introduced the crowding index $\alpha_i^c(\epsilon)$ defined by

$$\alpha_i^c(\epsilon) = \frac{\log p_i(\epsilon)}{\log \epsilon} \tag{A.4}$$

and only finite values of ϵ. Further, it is assumed that the cubes are locally isotropic. In the limit $\epsilon \quad 0$, α_i^c becomes α_i.

A limitation of this box-counting approach and fractal dimension in general, is that it is based on weak mathematical concepts. Pointwise dimensions of multifractals can not be calculated for finite point distributions. In addition, it is difficult to calculate fractal dimensions and various estimations reported different results. The application and interpretation for astrophysical applications and more discussions are found in Atmanspacher et al. (1989), Colombi et al. (1992), and Borgani (1995).

A.2 Details about Mock Catalogues

The mock catalogs, the numerical simulation of the observed large-scale galaxy distribution, as used in in Section 3.3 were created by Cole et al. (1998). For our analysis, we used their cosmological model L3, which is based on the cold dark matter scenario (CDM) and uses a functional form of the matter power spectrum $P(k) \quad f(q)k^n$ given by Bardeen et al. (1986), where $q = k\ \Gamma$ and $k = 2\pi\ \lambda$ is the wavenumber in units of $h\,\mathrm{Mpc}^{-1}$. The index n is the slope of the primordial power spectrum, $n = 1$, as predicted in the simplest models of inflation. The description of the power spectrum is completed by two further parameters, the shape parameter $\Gamma = 0\ 172$ and the amplitude $\sigma_8 = 1\ 05$, the linear theory root-mean-square fluctuations of the mass contained in spheres of radius $8\ h^{-1}$Mpc . These primordial density fluctuations evolve in the background cosmological model specified by the density parameter $\Omega_0 = 0\ 3$ and the cosmological constant $\Lambda = 0\ 7$. The baryon fraction was fixed using the constraint from primordial nucleosynthesis, $\Omega_\mathrm{b} = 0\ 0125\,h^{-2}$ (Walker et al. 1991), where a value of $h = 0\ 65$ was chosen.

A.3 Additional Plots: SIM Results for Galaxy Sample

In order to illustrate the meaning of geometry and the scaling index α, a example region which was reduced to two spatial dimensions was marked with the individual α-values (see also Section 3.4). A thin slice with a thickness of $4\,h^{-1}$Mpc was extracted from the volume limited SDSS galaxy sample and projected to a plane. Each row corresponds to a certain α-range and each galaxy was marked with symbols if its α-value falls into this range marked in the histogram on the right.

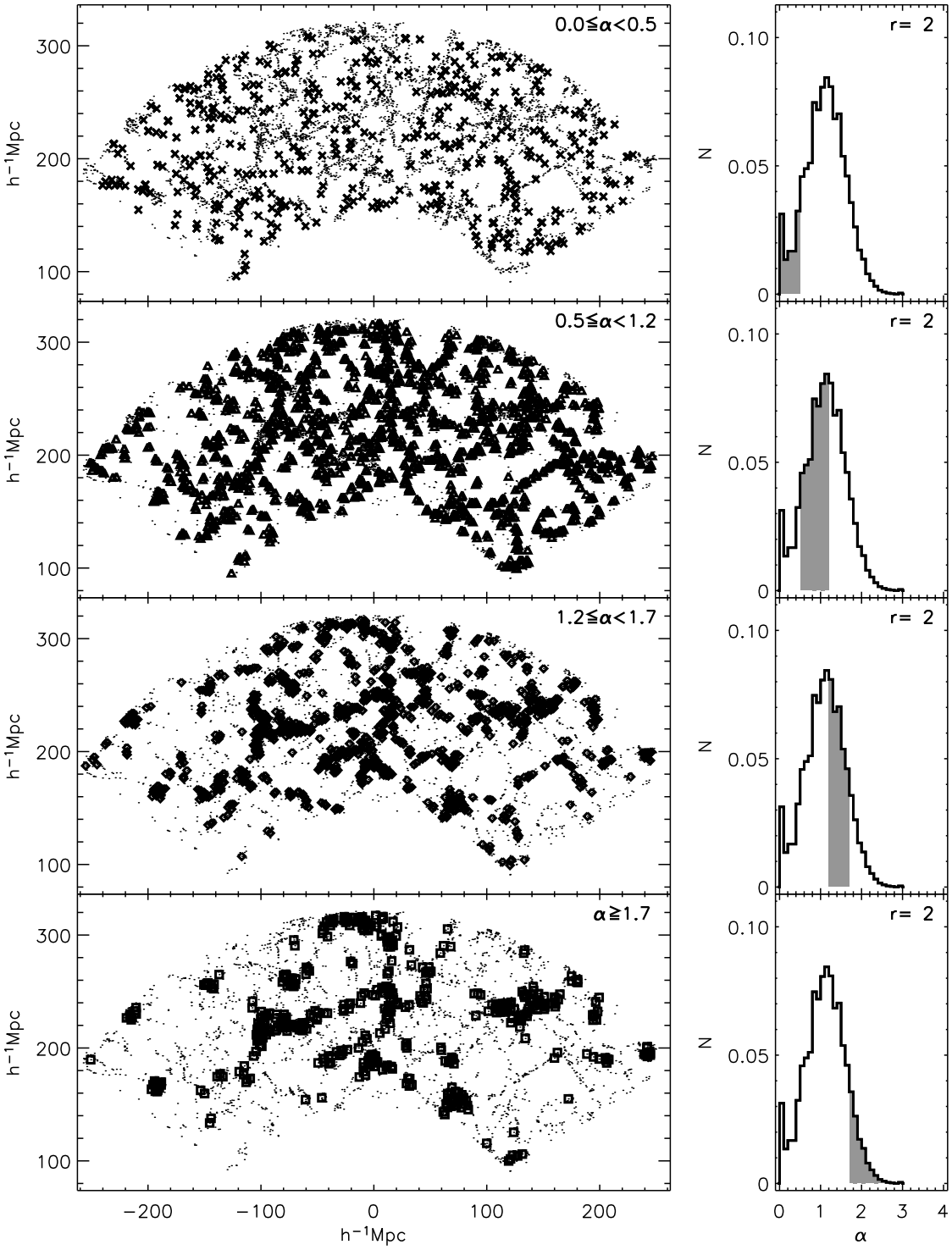

Figure A.1: Galaxy distribution encoded with different α ranges for the radius $r = 2\,h^{-1}$Mpc

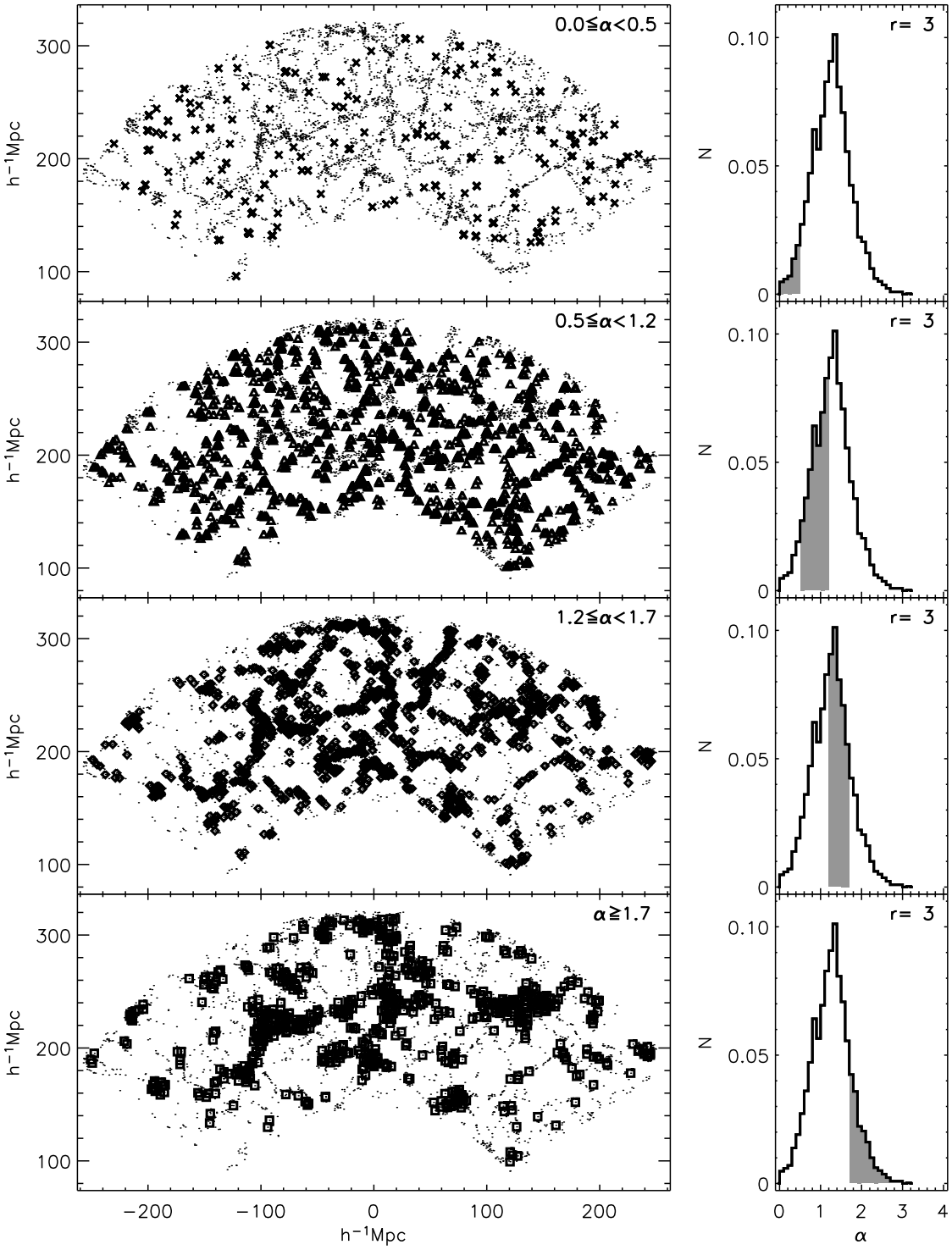

Figure A.2: Galaxy distribution encoded with different α ranges for the radius $r = 3\,h^{-1}$Mpc .

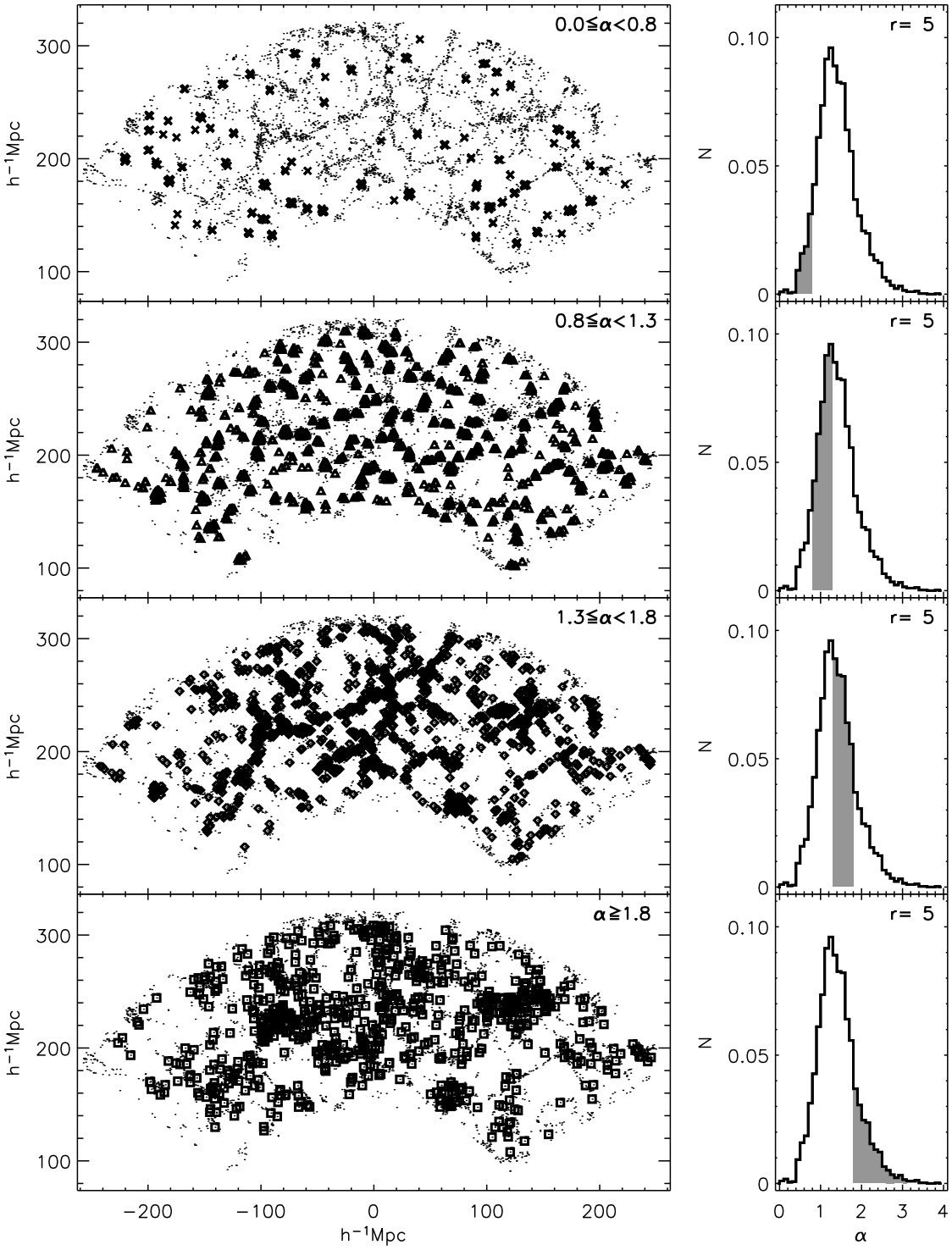

Figure A.3: Galaxy distribution encoded with different α ranges for the radius $r = 5\,h^{-1}$Mpc .

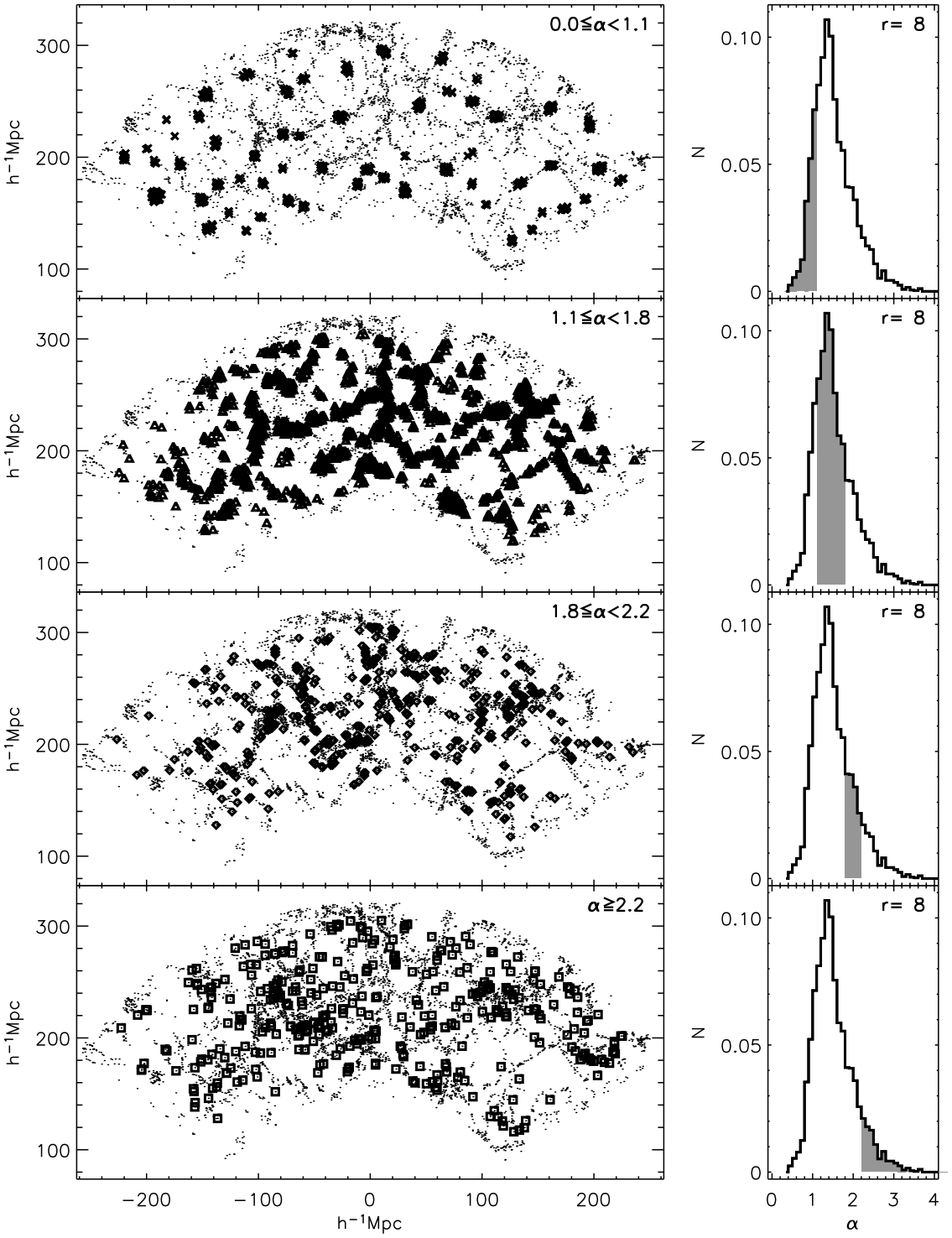

Figure A.4: Galaxy distribution encoded with different α rangesfor the radius $r = 8\,h^{-1}$Mpc .

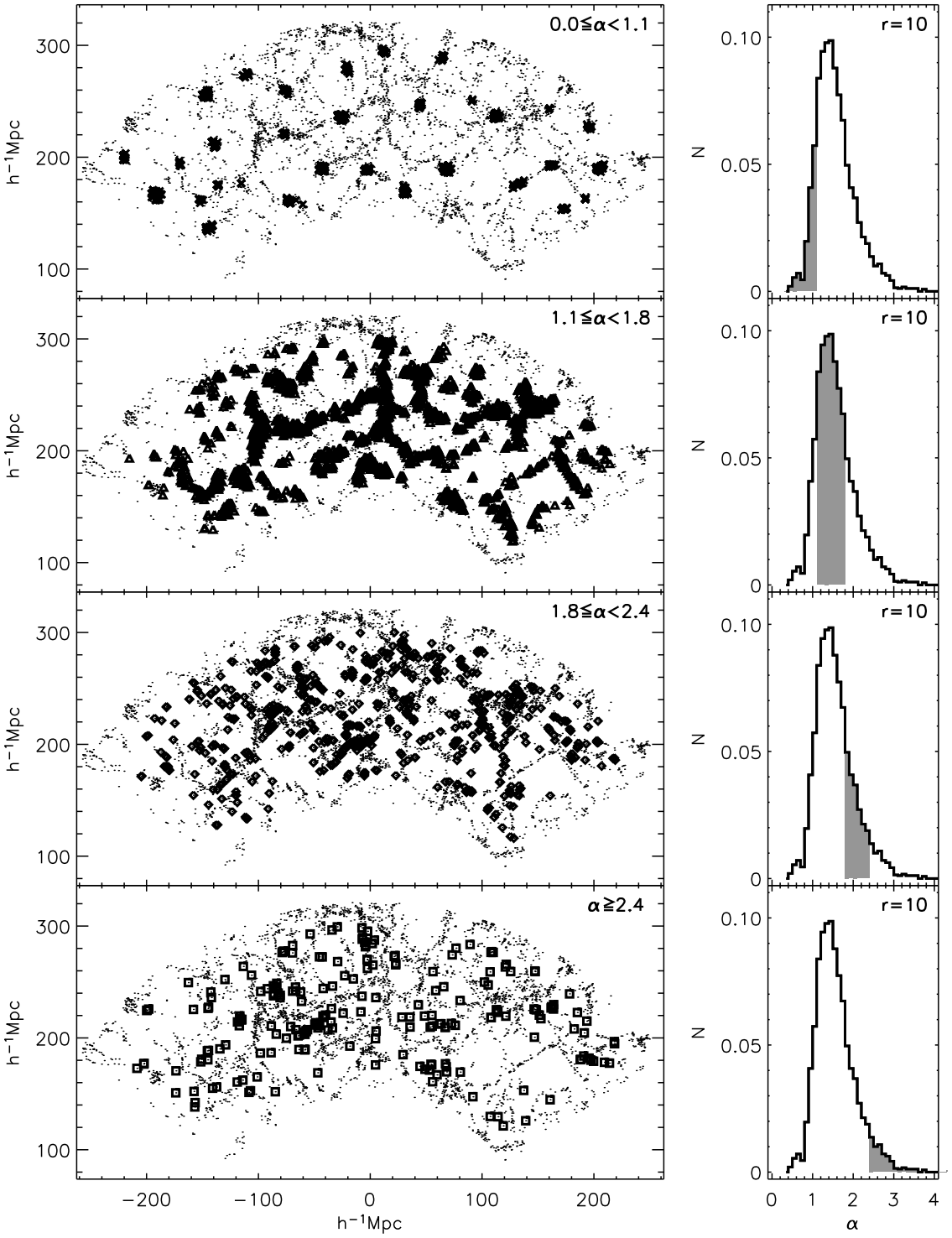

Figure A.5: Galaxy distribution encoded with different α ranges for the radius $r = 10\,h^{-1}$Mpc .

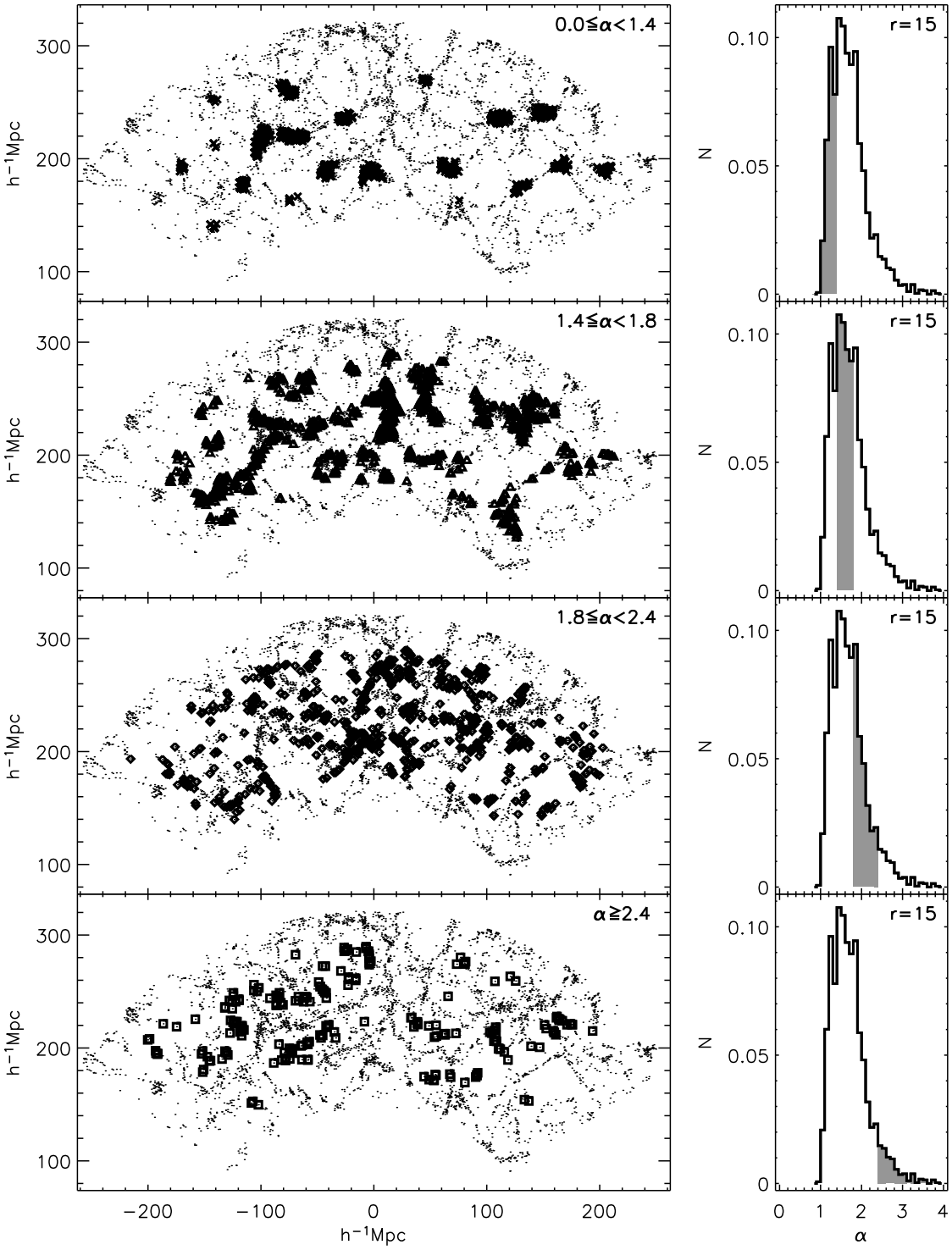

Figure A.6: Galaxy distribution encoded with different α ranges for the radius $r = 15\,h^{-1}$Mpc .

Appendix B

Source Code

B.1 SIM Source Code

The C++ source code is presented for calculation of the scaling index α including a kD-tree implementation to find close neighbor galaxies:

```
#include "kdtree2.hpp"
#include <boost/multi_array.hpp>
#include <vector>
#include <iostream>
#include <fstream>

using namespace boost;
using namespace std;

// define, for convenience a 2d array of floats.
typedef multi_array<float,2> array2dfloat;

int main() {

  //*** read data file
  long n_ele,n_columns,d,i,j,ir;
  //file with particles for which density is calculated
  char *data="/home/mhuber/Software/kdtree2.24/src-c++/input.ib";
  char *output="/home/mhuber/Software/kdtree2.24/src-c++/output.ib";

  const int n_rad=6;//number of different radii
  float const r[n_rad]={2.,3.,5.,8.,10.,15.};//radius r_0
  float const qp=2.;//exponent
  float r2[n_rad];
  for(i=0;i<n_rad;i++) r2[i]=pow(r[i],qp);// for later calulation
  float const treshold=pow(4.*r[n_rad-1],2);     // (4*max_radius)^2
  float dummy2;  //dummy variables
  vector<float> arri(n_rad),abaj(n_rad); //numerator and denominator for later calc.

  //*** load data file
  ifstream chunk(data);
  chunk.read( (char *) &n_columns, sizeof n_columns);   // all particles
  chunk.read( (char *) &n_ele, sizeof n_ele);   // all particles
```

```
 cout << "number of all particles: " << n_ele <<" columns: " <<n_columns<< endl;

 float * x = new float[n_ele];
 chunk.read( (char *) x, n_ele*4);
 float * y = new float[n_ele];
 chunk.read( (char *) y, n_ele*4);
 float * z = new float[n_ele];
 chunk.read( (char *) z, n_ele*4);
 chunk.close();

//*** prepare for search;
vector<float> qv(3);

float **alpha= new float *[n_rad]; //initialize alpha matrix
for(ir=0;ir<n_rad;ir++) alpha[ir]= new float[n_ele];
for(ir=0;ir<n_rad;ir++)
  for(j=0;j<n_ele;j++) alpha[ir][j]=0.;

kdtree2* tree;
kdtree2_result_vector results;
array2dfloat realdata;
realdata.resize(extents[n_ele][3]);
d=3;
for (i=0; i<n_ele; i++) {
  realdata[i][0] = x[i];
  realdata[i][1] = y[i];
  realdata[i][2] = z[i];
}

//*** create tree and rearrange
tree = new kdtree2(realdata,true);
tree->sort_results = true;         //sort results

//*** calculate alphas
for(i=0;i<n_ele;i++){
  for(j=0;j<d;j++)
    qv[j]=realdata[i][j];
  tree->r_nearest(qv,treshold,results);
  for(ir=0;ir<n_rad;ir++){
    arri[ir]=0.;
    abaj[ir]=0.;
  }
  for (unsigned int k=0; k<results.size(); k++) {
    for(ir=0;ir<n_rad;ir++){
  dummy2 =results[k].dis/r2[ir];
  arri[ir]+=dummy2*exp(-dummy2);
  abaj[ir]+=exp(-dummy2);
}
  }//k loop
  for(ir=0;ir<n_rad;ir++)
    alpha[ir][i]= (qp*arri[ir])/abaj[ir];
  if((i+1)%50000==0) cout<<round(100.*(i-1)/n_ele)<<"% done.. "<<endl;
}//i loop
```

```
 delete tree;

 //*** write results to file
 ofstream fout(output);
 fout.write( (char *) &n_rad, sizeof n_rad);
 fout.write( (char *) &n_ele, sizeof n_ele);
 for(ir=0;ir<n_rad;ir++) fout.write( (char *) alpha[ir], n_ele*4 );
 fout.close();
 delete [] x; delete [] y;delete [] z;
 return(0);
}
```

B.2 SDSS Queries

The photometric and spectroscopic data for the galaxies was extracted with the SQL database tool `sdssQA`. The following definitions were used to extract the data:

```
SELECT
-- object IDs
G.objID,
-- coordinates
S.ra,S.dec,S.z,
--modelMag_u,modelMag_g,modelMag_r,modelMag_i,modelMag_z,
G.dered_u,G.dered_g,G.dered_r,G.dered_i,G.dered_z,
G.modelMagErr_u,G.modelMagErr_g,G.modelMagErr_r,
G.modelMagErr_i,G.modelMagErr_z,

--extinction_u,extinction_g,extinction_r,extinction_i,extinction_z
G.petroMag_r,G.petroMagErr_r,G.extinction_r,
G.petroR50_r,G.petroR50Err_r,G.petroR90_r,G.petroR90Err_r

-- spectral data:
-- H_alpha
L1.ew,L1.ewErr,L1.continuum,
-- H_beta
L2.ew,L2.ewErr,L2.continuum,
-- OII
L3.ew,L3.ewErr,L3.continuum,
-- OIII
L4.ew,L4.ewErr,L4.continuum,
-- OIII 2
L5.ew,L5.ewErr,L5.continuum

FROM Galaxy G, SpecObj S, SpecLine L1, SpecLine L2,
     SpecLine L3,SpecLine L4, SpecLine L5

WHERE
S.bestObjID=G.ObjID
and (G.primTarget & 0x00000040) > 0 -- main galaxy sample
and S.specClass=2          -- spetro. galaxy

and S.specObjID=L1.specObjID -- connect two tables
```

```
and S.specObjID=L2.specObjID -- connect two tables
and S.specObjID=L3.specObjID -- connect two tables
and S.specObjID=L4.specObjID -- connect two tables
and S.specObjID=L5.specObjID -- connect two tables

and L1.lineID=6565               -- Ha line
and L2.lineID=4863               -- Hb line
and L3.lineID=3727               -- OII line
and L4.lineID=4960               -- OIII line
and L5.lineID=5008               -- OIII 2 line
```

Bibliography

Abadi, M. G., Moore, B., & Bower R. G. 1999, MNRAS, 308, 947

Abazajian, K., et al. 2003, AJ, 126, 2081

Adelman-McCarthy, J. K., et al. 2006, ApJS, 162, 38

Akritas, M. G., & Bershady, M. A. 1996, ApJ, 470, 706

Atmanspacher, H., Scheingraber, H., & Wiedenmann, G. 1989, Phys. Rev. A, 40, 3954

Baldry, I. K., Glazebrook, K., Brinkmann, J., Ivezić, Ž., Lupton, R. H., Nichol, R. C., & Szalay, A. S. 2004, ApJ, 600, 681

Balogh, M. L., Navarro, J. F., & Morris, S. L. 2000, ApJ, 540, 113

Balogh, M. L., et al. 2004, MNRAS, 348, 1355

Balogh, M. L., Baldry, I. K., Nichol, R., Miller, C., Bower, R., & Glazebrook, K. 2004, ApJ, 615, L101

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, A. S. 1986, ApJ, 304, 15

Barnes, J. E. & Hernquist, L. E. 1991, ApJ, 370, L65-L68

Barrow, J. D., Bhavsar, S. P., & Sonoda, D. H. 1984, MNRAS, 210, 19

Barrow, J. D., Bhavsar, S. P., & Sonoda, D. H. 1985, MNRAS, 216, 17

Bekki, K., Couch, W. J., & Shioya, Y. 2002, ApJ, 577, 651

Benson, A. J., Frenk, C. S., Baugh, C. M., Cole, S., & Lacey, C. G. 2003, MNRAS, 343, 679

Benson, A. J., Bower, R. G., Frenk, C. S., Lacey, C. G., Baugh, C. M., & Cole, S. 2003, ApJ, 599, 38

Berlind, A. A., Blanton, M. R., Hogg, D. W., Weinberg, D. H., Davé, R., Eisenstein, D. J., & Katz, N. 2005, ApJ, 629, 625

Berlind, A. A., Kazin, E., Blanton, M. R., Pueblas, S., Scoccimarro, R., Hogg, D. W. 2006, arXiv:astro-ph/0610524

Bertschinger, E., Dekel, A., Faber, S. M., Dressler, A., & Burstein, D., ApJ, 364, 370

Best, J. S., Charlton, J. C., & Mayer-Kress, G. 1996, ApJ, 456, 55

Bhavsar, S. P., & Ling, E. N. 1988, ApJL, 331, L63

Binney, J., & Tremaine, S. 2008, Galactic Dynamics: Second Edition, by James Binney and Scott Tremaine. ISBN 978-0-691-13026-2 (HB). Published by Princeton University Press, Princeton, NJ USA, 2008.,

Braglia, F., Pierini, D., Bohringer, H. 2007, A&A, 470, 425

Blanton, M. R., et al. 2001, ApJ, 121, 2358

Blanton, M. R., et al. 2003a, ApJ, 125, 2348

Blanton, M. R., et al. 2003b, ApJ, 594, 186

Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734

Blanton, M. R., & Moustakas, J. 2009, ARA&A, 47, 159

Borgani, S. 1995, Physics Reports, 251, 1

Bower, R. G., & Balogh, M. L. 2004, Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution, 325

Brinchmann, J., Charlot, S., White, S. D. M., Tremonti, C., Kauffmann, G., Heckman, T. & Brinkmann, J., 2004, MNRAS, 351, 1151

Brosche, P. 1973, A&A, 23, 259

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000

Butcher, H. & Oemler A., 1978, ApJ, 219, 18

Butcher, H. & Oemler A., 1984, ApJ, 285, 426

Byrd, G. & Valtonen M., 1990, ApJ, 350, 89

Carlberg, R. G., Yee, H. K. C., Ellingson, E., et al. 1996, ApJ, 462, 32

Cardelli, J. A., Clayton, G. C. & Mathis, J. S., ApJ, 345, 245

Charlot, R.G. & Longhetti, M., 2001, MNRAS, 323, 887
Christlein, D., & Zabludoff, A. I. 2005, ApJ, 621, 201
Colberg, J. M., Krughoff, K. S., & Connolly, A. J. 2005, MNRAS, 359, 272
Cole, S., Hatton, S., Weinberg, D. H., & Frenk, C. F. 1998, MNRAS, 300, 945
Cole, S., Lacey, C. C., Baugh, C. M., & Frenk, C. F. 2000, MNRAS, 319, 168
Cole, S., et al. 2005, MNRAS, 362, 505
Coles, P., & Chiang, L.-Y. 2000, Nature, 406, 376
Colless, M., et al. 2003, arXiv:astro-ph/0306581
Colombi, S., Bouchet, F. R., & Schaeffer, R. 1992, A&A, 263, 1
Couch, W. J., Barger, A. J., Smail, I., Ellis, R. S., & Sharples R. M. 1998, ApJ, 497, 188
Cowie, L., & McKnee, C. F. 1977, ApJ, 211,135
Cowie, L., Songalia, A., Hu, E. M., & Cohen, J. G. 1996, AJ, 112, 3
Cressie, N. A. C. *Statistics for Spatial Data*, John Wiley & Sons, New York, 1993
Davis, M., & Peebles, P. J. E. 1983, ApJ, 267, 465
Deeming, T. J. 1964, MNRAS, 127, 493
Dekel, A., & Lahav, O. 1999, ApJ, 520, 24
De Propris, R., et al. 2003, MNRAS, 342, 725
Diaferio, A., Kauffmann, G. , Balogh, M. L., White S. D. M., Schade, D. & Ellingson, E. 2001, MNRAS, 323, 999
Domínguez, M., Muriel, H., & Lambas, D. G. 2001, ApJ, 121, 1266
Doroshkevich, A. G., Muller, V., Retzlaff, J., & Turchaninov, V. 1999, MNRAS, 306, 575
Doroshkevich, A. G., Tucker, D. L., Fong, R., Turchaninov, V., & Lin, H. 2001, MNRAS, 322, 369
Doroshkevich, A., Tucker, D. L., Allam, S., & Way, M. J. 2004, A&A, 418, 7
Dressler, A. 1980, ApJ, 236, 351
Dressler, A. 1997, ApJ, 490, 577
Dressler, A., Smail, I., Poggianti, B. M., Butcher, H., Couch, W. J., Ellis, R. S., & Oemler, A., Jr. 1999, ApJS, 122, 51
Efron, B., 1979, Ann. Stat., 7, 1
Efron, B., & Tibshirani, R. J. 1993, *An Introduction to the Bootstrap*, Chapman & Hall, New York
Efstathiou, G., & Fall, S. M. 1984, MNRAS, 206, 453
Einasto, J., Joeveer, M., & Saar, E. 1980, MNRAS, 193, 353
Einasto, M., & Einasto, J. 1987, MNRAS, 226, 543
Einasto, M., Einasto, J., Muller, V., Heinamaki, P., & Tucker, D. L. 2003, A&A, 401, 851
Einasto, J., Tago, E., Einasto, M., et al. 2005, A&A, 439, 45
Eisenstein, D. J., et al. 2001, AJ, 122, 2267
Eisenstein, D. J., et al. 2005, ApJ, 633, 560
Evrard, A. E., Silk, J., & Szalay, A. S. 1990, ApJ, 365, 13
Folkes, S., et al. 1999, MNRAS, 308, 459
Fry, J. N., & Gaztanaga, E. 1994, ApJ, 425, 1
Fry, J. N., & Peebles, P. J. E. 1978, ApJ, 221, 19
Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748
Gnedin, O. Y. 2003, ApJ, 589, 752
Ghigna, S., Borgani, S., Bonometto, S. A., Guzzo, L., Klypin, A., Primack, J. R., Giovanelli, R., & Haynes, M. P. 1994, ApJL, 437, L71
Gómez, P., et al. 2003, ApJ, 584, 210
Goto T., et al. 2003a, PASJ, 55, 757
Goto T., et al. 2003b, ApJ, 346, 601
Gott, J. R., III, & Thuan, T. X. 1976, ApJ, 204, 649
Gott, J. R., III, Dickinson, M., & Melott, A. L. 1986, ApJ, 306, 341
Gott, J. R. I., et al. 2008, ApJ, 675, 16
Grassberger, P., & Procaccia, I. 1983, Physical Review Letters, 50, 346
Grassberger P., Badii R.,& Politi A. 1988, Journ. of Stat. Phys., 51, 135
Groth, E. J., & Peebles, P. J. E. 1977, ApJ, 217, 385
Gunn, J. E., et al. 1998, AJ, 116, 3040

Gunn, J. E., & Gott, J. R. 1972, ApJ, 176, 1

Hamilton, A. J. S. 1998, in The Evolving Universe, ed. D. Hamilton (Dordrecht: Kluwer), 185

Hawking, S. W., & Ellis, G. F. R. 1973, *The large-scale structure of space-time*, Cambridge (UK): Cambridge University Press

Hikage, C., et al. 2003, PASJ, 55, 911

Hashimoto, Y., Oemler, A. J., Lin, H., & Tucker, D. L. 1998, ApJ, 499, 589

Hogg, D. W. 1999, `astro-ph/9905116`

Hogg, D. W., Finkbeiner, D. P., Schlegel, D. J., & Gunn, J. E. 2001, AJ, 122, 2129

Hogg, D. W., Baldry, I. K., Blanton, M. R., & Eisenstein, D. J., 2002, `astro-ph/0210394`

Hogg, D. W., et al. 2003, ApJ, 585, L5

Hopkins, A. M., Connolly, A. J., Haarsma, D. B., & Cram, L. E., 2001, ApJ, 122, 288

Hopkins, A. M., Miller, C. J., Nichol, R. C., Connolly, A. J., Bernardi, M., Gómez, P. L., Goto, T., Tremonti, C. A., Brinkmann, J., Ivezić, Ž., Lamb, D. Q., 2003, ApJ, 599, 971

Hopkins, A. M., 2005, `astro-ph/04071170`

Hoyle, F., Rojas, R. R., Vogeley, M. S., & Brinkmann, J., 2005, ApJ, 620, 618

Huber, M. B., Diploma thesis, LMU (in German)

Huber, M. B., Carballido-Gamio, J., Fritscher, K., Schubert, R., Haenni, M., Hengg C., Majumdar, S., Link, T. M. 2009, Medical Physics, 36, 11

Hubble, E. 1926, ApJ, 64, 321

Hubble, E. & Humason, M. L., 1931, ApJ, 74, 43

Hubble, E. P. 1936, Realm of the Nebulae, by E.P. Hubble. New Haven: Yale University Press, 1936.

Huertas-Company, M., Foex, G., Soucail, G., & Pelló, R. 2009, A&A, 505, 83

Humason, M. L. 1936, ApJ, 83, 10

Isobe, T., Feigelson, E. D., Akritas, M. G., & Babu, G. J. 1990, ApJ, 364, 104

James, J. B., Colless, M., Lewis, G. F., & Peacock, J. A. 2009, MNRAS, 394, 454

Jamitzky, F., Stark, W., Bunk, W., Thalhammer, S., Raeth, C., Aschenbrenner, T., Morfill, G., & Heckl, W. 2001, Ultramicroscopy, 86, 241

Jeans, J. H. 1902, Royal Society of London Philosophical Transactions Series A, 199, 1

Jones, D. H., et al. 2009, arXiv:0903.545

Kaiser, N. 1984, ApJL

Kaiser, N. 1987, MNRAS, 227, 1

Kaiser, N., & Peacock, J. A. 2003, ApJ, 379, 482

Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, MNRAS, 264, 201

Kauffmann, G., et al. 2003, MNRAS, 341, 33

Kauffmann, G., et al. 2003, MNRAS, 341, 54

Kauffmann, G., et al. 2004, MNRAS, 353, 713

Kirshner, R. P., Oemler, A. J., Schechter, P. L., & Shectman, S. A. 1987, ApJ, 314, 493

Kennel, M. B. 2004, `physics/0408067`

Kennicutt, R. C. 1998, ARA&A, 36, 189

Kennicutt, R. C. 1998, ApJ, 498, 541

Kerscher, M., et al. 1997, MNRAS, 284, 73

Kerscher, M., Schmalzing, J., Buchert, T., & Wagner, H. 1998, A&A, 333, 1

Kerscher, M. 1999, A&A, 343, 333

Kerscher, M., et al. 2001, A&A, 377, 1

Knebe, A., Muller, V. 2000, A&A, 354, 761

Komatsu, E., et al. 2010, arXiv:1001.4538

Klypin, A., & Shandarin, S. F. 1993, ApJ, 413, 48

Kruskal J. B. 1956, Proc. Am. Math. Soc., 7(1), 48

Krzewina, L. G., & Saslaw, W. C. 1996, MNRAS, 278, 869

Landy, S. D., & Szalay, A. S. 1993, ApJ, 412, 64

Larson, R. B., Tinsley, B. M., & Caldwell, C. N. 1980, ApJ, 237, 692

Lehnert, M. D., et al. 2010, arXiv:1010.4312

Lynden-Bell, D., Faber, S. M., Burstein, D., Davies, R. L., Dressler, A., Terlevich, R. J., & Wegner, G. 1988, ApJ, 326, 19

Lewis, I., et al. 2002, MNRAS, 334, 673
Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999, AJ, 118, 1406
Madau, P., Ferguson, H. C., Dickinson, M. E., Giavalisco, M., Steidel, C. C., & Fruchter, A. 1996, MNRAS, 283, 1388
Madau, P., Pozzetti, L., & Dickinson, M. 1998, ApJ, 498, 106
Madgwick, D. S., et al. 2002, MNRAS, 333, 133
Mandelbrot, B. B. 1982, *The Fractal Geometry of Nature*, San Francisco: Freeman
Martinez, V. J., Paredes, S., & Saar, E. 1993, MNRAS, 260, 365
Martinez, V., & Saar, E. 2002, *Statistics of the Galaxy Distribution*, Chapman & Hall/CRC, Boca Raton
Martínez, H. J., Zandivarez, A., Merchán, M. E., & Domínguez, M. J. L. 2002, MNRAS, 337, 1441
Mecke, K. R., Buchert, T., & Wagner, H. 1994, A&A, 288, 697
Mihos, J. C., & Hernquist L. 2002, ApJ, 464, 641
Miller, C. J., Nichol, R. C., Gómez, P. L., Hopkins, A. M.,& Bernardi, M. 2003, ApJ, 597, 142
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Moore, B., Katz, N., Lake, G., Dressler, A., & Oemler A. 1996, Nature, 379, 613
Moore, B., Lake, G. & Katz, N. 1998, ApJ, 495, 139
Moss, C., & Whittle, M. 2000, MNRAS, 317, 667
Morgan, W. W. 1958, PASP, 70, 364
Mulchaey, J. S., & Zabludoff 1999, ApJ, 514, 133
Murtagh, F., & Heck, A. 1987, *Multivariate Data Analysis*, Reidel, Dordrecht
Nakamura, O., Masataka, F., Yasuda, N., Loveday, J., Brinkmann, J., Schneider, D. P., Shimasaku, K., & SubbaRao, M. 2003, ApJ, 125, 1688
Nikolic, B., Cullen, H., & Alexander, P. 2004, MNRAS, 355, 874
Oke, J. B., & Sandage, A. 1968, ApJ, 154, 21
Oemler, A. J. 1974, ApJ, 194, 1
Pandey, B., & Bharadwaj, S. 2006, MNRAS, 372, 827
Park, C., & Choi, Y.-Y. 2009, ApJ, 691, 1828
Park, C., & Hwang, H. S. 2009, ApJ, 699, 1595
Peacock, J. A. 1999, *Cosmological Physics*, Cambridge University Press
Peebles, P. J. E. 1973, ApJ, 185, 413
Peebles, P. J. E. 1980, *The Large-scale structure of the Universe*, Princeton Univesity Press 185, 413
Peebles, P. J. E. 1993, *Principles of Physical Cosmology*, Princeton Univesity Press
Peebles, P. J. E., & Nusser, A. 2010, Nature 465, 565
Petrosian, V. 1976, ApJ, 209, L1
Pietronero, L. 1987, Physica A , 144, 257
Pimbblet, K. A., Drinkwater, M. J., & Hawkrigg, M. C. 2004, MNRAS, 354, L61
Postman, M., & Geller, M. J. 1984, ApJ, 281, 95
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Quilis, V., Moore, B., & Bower, R. 2000, Science, 288, 1617
Raeth, C., Bunk, W., Huber, M. B., Morfill, G., Retzlaff, J., Schuecker, P. 2002, MNRAS, 337, 413
Raeth, C., and Schuecker, P. 2003, MNRAS, 344, 115
Rath, C., Schuecker, P., & Banday, A. J. 2007, MNRAS, 380, 466
Richards, G. T., et al. 2002, AJ, 123, 2945
Rojas, R. R., Vogeley M. S., Hoyle, F., & Brinkmann, J. 2004, ApJ, 617, 50
Rojas, R. R., Vogeley M. S., Hoyle, F., & Brinkmann, J. 2005, ApJ, 624, 571
Rossmanith, G., Raeth, C., Banday, A. J., & Morfill, G. 2009, arXiv:0905.2854
Salpeter, E. E. 1955,ApJ, 121, 161
Schmalzing, J., & Diaferio, A. 2000, MNRAS, 312, 638
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998,ApJ, 500, 525
Scoccimarro, R., Colombi, S., Fry, J. N., Frieman, J. A., Hivon, E., & Melott, A. 1998, ApJ, 496, 586
Scoccimarro, R., Feldman, H. A., Fry, J. N., & Frieman, J. A. 2001, ApJ, 546, 652
Seljak, U., et al. 2005, Physical Review D, 71, 043511
Sérsic, J. L. 1968, Cordoba, Argentina: Observatorio Astronomico, 1968
Silverman, B. W. 1986, *Density Estimation for Statistics and Data Analysis*, Chapman & Hall

Shimasaku, K., et al. 2001,ApJ, 122, 1238
Smith, J. A., et al. 2000, Bulletin of the American Astronomical Society, 32, 1424
Smith, G. P., Treu, T., Ellis, R. S., Moran, S. M., & Dressler, A. 2005, ApJ, 620, 78
Spergel, D. N., et al. 2003, ApJ, 148, 175
Spergel, D. N., et al. 2007, ApJS, 170, 377
Springel, V., et al. 2005, Nature, 435, 629
Stoughton, C., et al. 2002,ApJ, 123, 485
Strateva, I., et al. 2001,ApJ, 122, 1861
Strauss, M. A., et al. 2002,ApJ, 124, 1810
Sullivan, M., Mobasher, B., Chan, B., Cram, L., Ellis, R., Treyer, M., & Hopkins, A. 2001, ApJ, 558, 72
Taruya, A., & Suto, Y. 2000, ApJ, 542, 559
Tanaka, M.,Goto, T., Okamura, S., Shimasaku, K., & Brinkmann, J. 2004, AJ, 128, 2677
Tegmark, M., et al. 2004a, Physical Review D, 69, 103501
Tegmark, M., et al. 2004b, ApJ, 606, 702
Tempel, E., Einasto, J., Einasto, M., Saar, E., & Tago, E. 2009, A&A, 495, 37
Toomre, A., and Toomre, J., 1972, ApJ, 178, 623
Totsuji, H., & Kihara, T., 1969, PASJ, 21, 221
Tremonti, C. A., et al. 2004, ApJ, 613, 898
Tucker, D. L., Oemler, A., Jr., Kirshner, R. P., et al. 1997, MNRAS, 285, L5
van den Bergh, S. 1976, ApJ, 206, 886
van de Weygaert, R. 1991, PhD thesis, Univ. Leiden
Walker, T. P., Steigman, G., Kang, H.-S., Schramm, D. M., & Olive, K. A. 1991, ApJ, 376, 51
Weinberg, D. H., Davé, R., Katz, N., & Hernquist, L. 2004, ApJ, 601, 1
White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341
White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52
White, S. D. M., Frenk, C. S., Davis, M., & Efstathiou, G. 1987, ApJ, 313, 505
White, S. D. M. 1979, MNRAS, 186, 145
Whitmore, B. C., & Gilmore, D. M. 1991, ApJ, 367, 64
Whitmore, B. C, Gilmore, D. M., & Jones, C. 1993, ApJ, 407, 489
Yoshida, N., et al. 2001, AJ, 325, 803
York, D., et al. 2000, AJ, 120, 1579
Zehavi, I., et al. 2002, ApJ, 571, 172
Zehavi, I., et al. 2004, `astro-ph/0408569`
Zeldovich, Y. B. 1970, A&A, 5, 84

Acknowledgements

I would like to thank my PhD advisors Professor Dr. Gregor Morfill and Dr. Christoph Rath for supporting me during all these years and providing me this opportunity. I would like to thank and remember Dr.habil. Peter Schuecker for his support and vision. Gregor, Christoph, and Peter guided and supported me during the Internship, Diploma thesis, and the PhD research at the MPE. I thank Gregor for giving me this unique research opportunity and this ongoing counsel and guidance along the way. I thank Christoph for his guidance, patience, and all his helpful suggestions for this work. I thank Peter for his scientific advice and all the insights I gained from him. I also would like to thank Professor Hans Bohringer for his support and helpful comments and for providing the access to the SDSS data.

I would like to thank my friends for providing support and friendship that I needed. They showed great interest in my work and provided many inspirations during all this time.

I thank my family for all their strong support and inspiration. I thank my mom and dad, Hermann, Yuka, and all my siblings for their continuous motivation, encouragement and understanding, and their interest in my work. I especially thank my sister Rebecca for reading the thesis and helping me with the submission.